JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 8 (2020), Issue 11 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Viewpoints

Original Papers

Reviews

Corrigenda and Addenda

Viewpoint

Opportunities for Mobile App–Based Adherence Support for Children With Tuberculosis in South Africa

Rachel M Morse[1], BSc; Hanlie Myburgh[2,3], MA; David Reubi[1], MA, MSc, PhD; Ava E Archey[4]; Leletu Busakwe[2,5], BA; Anthony J Garcia-Prats[2], MD, PhD; Anneke C Hesseling[2], MBChB, PhD; Stephanie Jacobs[2], BA; Sharon Mbaba[2], NDip; Kyla Meyerson[2], MA; James A Seddon[2,6], MBBS, MA, PhD, FRCPCH; Marieke M van der Zalm[2], MD, MSc, PhD; Dillon T Wademan[2], MA; Graeme Hoddinott[2], MSocSc, PhD

[1]Department of Global Health and Social Medicine, King's College London, London, United Kingdom

[2]Desmond Tutu Tuberculosis Centre, Department of Paediatrics and Child Health, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

[3]Amsterdam Institute for Social Science Research, University of Amsterdam, Amsterdam, Netherlands

[4]College of Arts and Sciences, University of Virginia, Charlottesville, VA, United States

[5]Social Aspects of Public Health, Human Sciences Research Council, Cape Town, South Africa

[6]Department of Infectious Diseases, Imperial College London, London, United Kingdom

Corresponding Author:
Hanlie Myburgh, MA
Desmond Tutu Tuberculosis Centre
Department of Paediatrics and Child Health
Faculty of Medicine and Health Sciences, Stellenbosch University
Lower Level Clinical Building
Francie van Zijl Drive
Cape Town, 7505
South Africa
Phone: 27 823416810
Email: hmyburgh@sun.ac.za

Abstract

Tuberculosis is the number one infectious cause of death globally. Young children, generally those younger than 5 years, are at the highest risk of progressing from tuberculosis infection to tuberculosis disease and of developing the most severe forms of tuberculosis. Most current tuberculosis drug formulations have poor acceptability among children and require consistent adherence for prolonged periods of time. These challenges complicate children's adherence to treatment and caregivers' daily administration of the drugs. Rapid developments in mobile technologies and apps present opportunities for using widely available technology to support national tuberculosis programs and patient treatment adherence. Pilot studies have demonstrated that mobile apps are a feasible and acceptable means of enhancing children's treatment adherence for other chronic conditions. Despite this, no mobile apps that aim to promote adherence to tuberculosis treatment have been developed for children. In this paper, we draw on our experiences carrying out research in clinical pediatric tuberculosis studies in South Africa. We present hypothetical scenarios of children's adherence to tuberculosis medication to suggest priorities for behavioral and educational strategies that a mobile app could incorporate to address some of the adherence support gaps faced by children diagnosed with tuberculosis. We argue that a mobile app has the potential to lessen some of the negative experiences that children associate with taking tuberculosis treatment and to facilitate a more positive treatment adherence experience for children and their caregivers.

(JMIR Mhealth Uhealth 2020;8(11):e19154) doi:10.2196/19154

KEYWORDS

eHealth; mHealth; tuberculosis; pediatric tuberculosis; adherence

Introduction

Tuberculosis (TB) is the number one infectious cause of death globally [1]. Implementation of effective TB control programs continues to place enormous pressure on the health systems of high TB-burden countries [1]. The World Health Organization's (WHO) Global TB Program identifies digital health technologies as key to supporting its new End TB Strategy [2]. An essential component of this strategy includes using mobile technologies and apps in particular to improve adherence to TB treatment [2]. Reviews show that the development and implementation of mobile health technologies to support health outcomes are overwhelmingly concentrated in high-income countries and focused on diseases of concern in those settings [3-5]. The development and implementation of mobile health technologies in low- and middle-income countries (LMICs), where the burden of communicable diseases, such as TB, is markedly high and treatment is strenuous, has received comparatively less attention. In the last decade, rapid developments in mobile technologies and apps, extended mobile network coverage, and increasing possibilities for integrating mobile health into existing eHealth platforms present opportunities for using widely available technology to support TB programs and adherence in resource-constrained and high-burden areas [2].

In South Africa, TB is the leading cause of death, mainly due to the overlapping HIV and TB epidemics [1,6]. South Africa also has the world's second highest estimated TB incidence rate, at 520 per 100,000 people [1]. In endemic areas, such as South Africa, young children (especially younger than 5 years) are at the highest risk to progress from TB infection to TB disease and to develop the most severe forms of TB, such as TB meningitis [7,8]. In 2018, children accounted for 7% of newly recorded TB cases in South Africa [1]. Despite their heightened risk, children's paucibacillary TB disease leads to them having much better treatment outcomes than their adult counterparts once treatment is started [7]. Children's timely diagnosis and rapid initiation onto TB treatment are critical, as undiagnosed cases or undertreatment could have long-term health consequences or lead to death.

Drug-susceptible (DS) and drug-resistant (DR) TB treatment formulations have had very poor acceptability among children [9]. The drugs are multiple, unpalatable, and often cause significant side effects, such as nausea, liver toxicity, insomnia [1,7], neurosensory alterations, neuromuscular weakness, and psychological and behavioral disturbances such as anger, aggression, withdrawal, hyperactivity, and depressive symptoms [10,11]. TB treatment is also lengthy and requires consistent adherence for at least six months for DS-TB and, in the case of DR-TB, adherence to different drug regimens for a longer period of time [1,8]. Child-friendly TB preventive therapy and treatment formulations that simplify and improve TB treatment for children are either only recently available or are still being evaluated in clinical trials [8,9,12]. The challenges presented above complicate children's adherence to treatment and caregivers' daily administration of the drugs [8].

Various small-scale and pilot studies in high-income settings have demonstrated that mobile apps are a feasible and acceptable means of enhancing children's and young people's treatment adherence for chronic conditions, such as sickle cell disease (SCD), migraines, asthma, and HIV [13-18]. There is limited high-quality evidence on the impact of mobile apps on health outcomes when implemented at scale or on how different mobile app components affect health and treatment outcomes [3-5]. Despite this, many of the small-scale studies published thus far provide promising evidence supporting the potential of mobile apps to positively impact adherence and treatment outcomes among children and young people [3-5,13,15,16,19]. These studies also serve as a foundation for further research and implementation efforts in this field and demonstrate the potential for developing and implementing mobile apps in LMICs that support children's and caregivers' burdensome TB treatment journeys. To date, no mobile apps that aim to improve the experience of TB and TB treatment in high- or low-income countries have been developed for children; mobile health interventions for supporting TB programs have largely focused on adult patients, using SMS-based text message platforms to promote TB treatment adherence, clinic attendance, and monitoring [20-22].

In this paper, we reflect on our joint experience conducting research in ongoing pediatric TB studies at the Desmond Tutu TB Centre, Stellenbosch University, South Africa, to highlight the potential for developing mobile app–based technologies to support TB treatment adherence in children. First, we provide an overview of technologies that have been used to support children's and young people's treatment and adherence experiences in other disease contexts. Second, we present hypothetical scenarios informed by our experiences to illustrate the adherence support gaps for children with TB. Third, we imagine how a mobile app could address some of the adherence support gaps faced by children affected by TB.

Overview of Adherence Technologies for Children

Table 1 provides examples of studies that use mobile technology to improve children's and young people's treatment adherence, including for SCD, asthma, migraine, attention-deficit/hyperactivity disorder, and HIV. These pilot studies provide preliminary evidence that shows the potential of mobile apps to support adherence in children and young people.

Table 1. Examples of studies that used mobile technology to improve children's and young people's treatment adherence.

Study	Study summary	Intervention features
Creary et al (2019) [13]	- Single-arm intervention study with pediatric patients ≤19 years (median age 10 years) - Condition: SCD[a] - Country: United States - Intervention: effect of mobile phone–based intervention on adherence to SCD treatment - Outcomes: increased adherence among participants engaged with the intervention over a 6-month period compared with participants' baseline adherence	- Self-recorded videos of taking the medication - Monetary incentives with successful adherence (US $30 gift card) - Text message reminders - Personalized feedback on adherence from research staff
Ramsey et al (2018) [14]	- AB design pilot study with patients aged 13 to 21 years - Condition: migraine - Country: United States - Intervention: evaluation of the use of a mobile app to increase adherence to migraine medication - Outcomes: similarly high rates of adherence among participants using the intervention and participants using smart pill bottles over an 8-week period. Participants rated the intervention as acceptable and easy to use	- Reminder notifications and phone calls - Incorporation of the health belief model - Medication adherence tracking
Kosse et al (2019) [15]	- Randomized controlled trial with patients aged 12 to 18 years - Condition: asthma - Country: Netherlands - Intervention: measuring the impact of a mobile app on participants' adherence to asthma treatment - Outcomes: increased 6-month adherence rates among participants who used the app and measured low-adherence rates at baseline. Decreased 6-month adherence rates measured among participants in the control group	- Symptom monitoring - Education (short movie clips) - Medication reminders - Chat feature with pharmacists - Chat feature with peers
Weisman et al (2018) [17]	- Randomized controlled trial with children aged 6 to 16 years - Condition: ADHD[b] - Country: Israel - Intervention: assessing the use of a mobile app to increase ADHD medication adherence - Outcomes: increased adherence rates among participants who were using the app over 8 weeks compared with adherence rates in the control group	- Symptom monitoring - Reminder notifications - Education on ADHD and its treatment (video clips, articles, and links to additional reading materials)
Curtis et al (2019) [23]	- Study conducted with 12- to 18-year-olds - Condition: SCD - Country: multiple, including Lebanon and Nigeria - Intervention: development of an evidence-based mobile app supporting SCD medication adherence - Outcomes: evaluation of the app is underway	- Avatar - Rewards system (points earned through interaction with app's features) - Education on SCD - Quiz (focused on education about SCD) - Symptom tracking - Medication and appointment reminders - Emergency information for health professionals unfamiliar with SCD
Hightow-Weidman et al (2018) [18]	- Study conducted with 19- to 24-year-olds - Condition: HIV - Country: United States - Intervention: pilot of a mobile app supporting HIV medication adherence - Outcomes: high acceptability and feasibility, as measured in participant ratings after 28 days of app use. A clinical trial evaluating adherence with the app is underway	- Medication tracking - Reward system (points and monetary incentive earned through interaction with the app's features) - Social networking (social wall discussion questions) - Gamification through quests and story-based framework - Incorporation of social cognitive theory

[a]SCD: sickle cell disease.

[b]ADHD: attention-deficit/hyperactivity disorder.

However, few technologies and mobile apps target children's treatment adherence specifically; for TB, no such technologies exist. Common app features shared across the studies included in Table 1 incorporated behavioral and educational strategies to improve adherence through features such as reminder notifications, adherence logging, cash or other adherence rewards, and educational features [13-15,17,18,23]. While evidence of the impact of these app components on adherence is limited to small sample sizes, the studies that have completed evaluations show increased or high rates of adherence in their study populations [13-16,19], and an even greater number of apps have shown high acceptability and feasibility among users from various age categories [13-18]. For example, one of the pilot studies, MyMate&Me, is a mobile app aiming to support SCD medication adherence among children and adolescents aged 12 to 18 years [23]. With MyMate&Me, users earn points by participating in app features, including (1) "Tip of the Day," where users learn important information about coping with SCD, (2) "Daily Quiz," where users learn about SCD by answering quiz questions and earning points for correct answers, (3) "Mood Tracker," where users earn points by logging their daily mood, and (4) "Medication and Appointment Reminders," where users are reminded to take their medication, attend appointments, and earn points for confirming their adherence and attendance. MyMate&Me also features an avatar to accompany the user throughout the app. The user can spend their earned points on clothes and accessories for their avatar, thereby providing incentive for using the app's various functionalities and managing their condition. Another study, also on SCD, incorporated video directly observed therapy for adherence monitoring, daily reminders via SMS text messages, email, or telephone if adherence was not observed, and a US $30 gift card incentive if users maintained at least 90% adherence over 30 days [13].

The behavioral strategies used in these mobile apps aimed to facilitate a favorable adherence environment and thus change patient behavior, while the educational strategies aimed to provide patients with information so that they could make informed choices about adhering to treatment [24]. These strategies are informed by underlying theories of human behavior, such as the health belief model [25], the theory of reasoned action or planned behavior [26], and operant conditioning [27]. The basic logic of these theories is that people are more likely to repeat behaviors (in this instance adherence) that (1) are associated with reward, not punishment, (2) conform to their perceptions of descriptive and injunctive norms, and (3) are within their perceived self-control. Robust evidence supporting the use of mobile apps and mobile app features that use behavioral strategies is limited [3-5]; however, their foundational principles have been evaluated extensively in relation to a large variety of health behaviors in many settings across the world [27,28] and have been shown to be supportive of adherence [24]. In 2014, the WHO advocated for such behavioral and educational strategies to be incorporated into apps to improve TB treatment adherence [7].

What Are the Current Adherence Support Gaps for Children With Tuberculosis?

Many mobile apps that support children's and young people's adherence to treatment for other conditions serve as treatment reminders or sources of information for patients [13-15,17,18,23]. Drug palatability, tolerability, and acceptability, which pose serious threats to adherence, are typically not considered for intervention with mobile apps. In Textbox 1, we draw on our experiences carrying out research in pediatric TB studies at the Desmond Tutu TB Centre to consider children's and caregivers' experiences with TB treatment that illustrate specific adherence support gaps and needs. The scenarios reflect typical interactions with patients and caregivers that are observed daily in TB service delivery contexts. They show that caregivers and children are generally motivated and knowledgeable about TB treatment but that the unpalatability, number, and side effects of the drugs affect their acceptability and provide serious challenges for caregivers and children at each ingestion event. We argue that an important user need is making the drug ingestion and administration experience more acceptable. For children especially, the current TB treatment experience is negative, and they are "punished" by the negative associations and poor palatability at each ingestion event. Alongside advocating for improved, child-friendly, palatable formulations, a mobile app that can mitigate the punishment associated with treatment and provide tangible rewards closely associated with each adherence event is a priority and could serve as a supporting component for improving the TB treatment experience for children and their caregivers. The app could also, for example, provide users with tutorial videos for pairing TB treatment with foods to improve palatability without compromising dosing accuracy [29].

Textbox 1. Scenarios reflecting specific adherence support gaps and needs.

Scenario 1

Bernice is the mother of 2-year-old Lindiwe and takes her role in administering Lindiwe's tuberculosis (TB) treatment very seriously. Every day when it is time to administer Lindiwe's treatment, Bernice endures frightful temper tantrums due to the drugs' poor taste. Bernice finds that the daily frustration and struggle of ensuring Lindiwe ingests her medication is mentally wearing, and she has had to be increasingly resourceful to plan for and strategize ways of getting Lindiwe to take her treatment, including begging, bribing, forcing, cheering, and more. In this scenario, the adherence gap is Bernice's struggle to be resourceful and find new ways to administer the TB medication due to the medication's poor acceptability.

Scenario 2

Angela's 7-year-old son, Henry, often experiences significant nausea when he takes his TB treatment. On occasion, the treatment causes him to vomit. The process of giving Henry his TB medication has become a daily battle for Angela and often involves her forcing the medication into him. This arrangement is beginning to negatively affect their relationship. The adherence support gap in this scenario is the negative effect of the TB medication's poor palatability and acceptability on Angela's relationship with Henry.

Scenario 3

Despite being only 5 years old, Alice understands that taking TB medication is an important daily task. When it is time for her mother, Gill, to take her own TB medication, Alice reminds her by saying "pills, Mommy, pills." However, understanding the importance of TB treatment does not help Alice's adherence; she puts up a great fight each time she has to take her own treatment. In this scenario, the adherence support gap is that Alice's understanding of the importance of TB treatment is not enough to overcome the poor acceptability of the treatment.

How Could Technology Address These Gaps?

In order to address the current treatment support gaps for children with TB in South Africa, we conceptualize a mobile app that incorporates behavioral and educational strategies to include an avatar, a rewards scheme, reminder notifications, adherence tracking, social support for children and caregivers, and information on TB, TB treatment, and TB treatment adherence, with the overall aim of reducing the negative emotions associated with TB treatment for children. The connections between treatment support gaps, strategies, and app features are presented in Table 2.

Table 2. Summary of adherence gaps and their hypothetical corresponding mobile app features.

Adherence gap	App features: educational	App features: behavioral
Poor acceptability and palatability of TB[a] treatment for children	Information about the rewards of adherence and the consequences of nonadherence	- An avatar to reduce children's negative emotions toward treatment and demonstrate the consequences of adherence - Rewards scheme to reduce children's negative emotions toward treatment - Social support for children from peers
Caregivers' challenges finding strategies to administer TB treatment	Access to advice and information from health professionals	- Reminder notifications and medication tracking associated with behavior-changing rewards - Social support and advice from other caregivers
Negative effect of the strain of TB treatment on the child-caregiver relationship	Access to advice and information from health professionals	- Rewards scheme to reduce children's negative emotions toward treatment - Social support and advice for caregivers from other caregivers - Social support for children from peers - An avatar to reduce children's negative emotions toward treatment
Poor acceptability of TB treatment outweighing the importance of TB treatment for children	Information about the rewards of adherence and the consequences of nonadherence	- An avatar to reduce children's negative emotions toward treatment and demonstrate the rewards of adherence - Rewards scheme to reduce children's negative emotions toward treatment - Social support for children from peers

[a]TB: tuberculosis.

We propose that the child creates an avatar to represent themselves within the app. The avatar is given opportunities to thrive (eg, growing noticeably stronger or getting better outfits or gear) by completing tasks so that patients are rewarded for active engagement in their care rather than punished for noncompliance, as is the current norm. The avatar serves as an externalization of the child's experiences with treatment. If the child cares for their avatar, then the avatar will grow increasingly sophisticated; achieving this is a reward in itself for the child. For example, in such an app, users may complete missions in order to gain diamonds—the currency of the app—that can be used to purchase clothes and accessories for the avatar in the app's store. Completing missions also leads to levelling up each time the user acquires the required number of diamonds. Each mission might include linked goals, such as following the treatment plan for 3 consecutive days, reading about TB, and joining an online forum discussion. As a user completes missions, earns diamonds, and levels up, their avatar becomes stronger. "Adherence streaks" unlock bonus rewards and items in the app's store: the longer the adherence streak, the greater the reward. Rewards can be scaled over time so that the user has increasingly higher targets to achieve. As a further incentive,

such an app could partner with businesses to offer small tangible rewards, such as mobile data or discounts at selected stores. The app, by encouraging the child's active participation in their treatment adherence, assigns positive consequences to taking their medication (eg, diamonds, a visibly stronger avatar, and more accessories) and permits the child to gain control and autonomy over the treatment experience. It also ties rewards to adherence despite the unpalatability of the drugs and negative side effects. Interventions that use rewards to promote medication adherence have been shown to be significantly more efficacious than interventions that do not use rewards [30]. The aforementioned features of the app's design also have the potential to lower the risk of user fatigue. In an app supporting TB treatment adherence, the risk of user fatigue is mitigated, as the need for using the app naturally ends upon treatment completion, at which point the user's avatar could enter a hall of fame.

Building on a fundamental link between the avatar's wellness and the child's adherence behaviors, a mobile app could then incorporate a range of other features to support children and caregivers in adhering to treatment. For instance, the app could be an important source of information on TB. It could include medication reminders, alerts for upcoming clinic visits, and details of what to expect at the visit. Daily adherence logging would allow users to track their adherence over time, view weekly and monthly adherence, and progress toward completing their treatment course. An in-app community space would allow caregivers to interact with other caregivers. Caregiver participation in the community would earn their child diamonds, making knowledge seeking collaborative and providing caregivers with a new adherence resource and a free, positive reward to help mitigate the negative effects of forced administration on the child-caregiver relationship. In the online forum, caregivers and older children with appropriate supervision would be able to ask each other about their experiences with TB and treatment, creating an additional social support system.

Studies of TB knowledge in South Africa have suggested relatively good levels of understanding of prevention and cure among the population, but knowledge of TB cause and symptoms was suggested to be low [31,32]. A frequently asked questions (FAQs) section, designed by children and adolescents who have experienced TB disease and medical professionals who work with TB disease, could address knowledge gaps by including answers to common questions, such as "How can I protect my family and friends from getting TB?" "What kind of bodily changes can I expect while taking treatment?" "Can I go to school if I have TB?" and "What is the relationship between TB and HIV?" In the app's community space, users can also play games (racing, sports, treasure hunts, etc), either with each other or solo, to earn additional diamonds. The stronger their avatar (eg, the better their adherence), the more competitively they will perform in the games. The effectiveness of such games for increasing medication adherence among children has been shown [33,34]. For example, in a randomized controlled trial, a video game intervention was found to improve medication adherence and behavioral outcomes for young people aged 13 to 29 years with acute leukemia, lymphoma, or soft-tissue sarcoma in the United States, Canada, and Australia [33].

In a high TB-burden, resource-limited setting like South Africa, such a mobile app that aims to facilitate children's adherence to TB treatment could improve the patient experience and also potentially buffer some of the shortcomings of the care provided in an overburdened health service. In particular, a mobile app could offer avenues of support to caregivers and children that are not immediately available in service encounters. It could also be a source of information on a range of topics related to the disease and its treatment, as in the FAQ section proposed. The immediate availability of such support and information has the potential to be empowering for patients.

Considerations for Implementation

There are a number of practical considerations for developing and using mobile apps to support children's adherence in culturally diverse, developing health settings, such as South Africa. First, most households in low socioeconomic settings have limited access to phones and tablets that could support apps. Programs may consider providing patients with smartphones during treatment, with the possibility of taking ownership when treatment has been completed successfully. The resources required to deliver such an app (if effective in improving adherence) are limited to the supply of a suitable mobile device (about US $40 per user) and some training on how to use the app. Studies suggest that the cost of providing smartphones to patients and supporting app-based adherence is more cost-effective for patients and health services than current adherence support models, which are often highly labor-intensive, requiring many hours of health workers' time [35,36]. The app would also require some tailoring to the age of the child to provide age-appropriate services and support, including features for older children and adolescents who may experience greater treatment autonomy. During the app's initial setup, a prompt to input age can align the app's functionality with the age of the child. To be acceptable amongst diverse users, the app must be attuned to the local context in which it will be implemented. As such, the app's avatar must offer an array of characteristics that users can select to reflect their social, cultural, racial, and gender identities and their personalities. These could be informed by participatory research through focus group discussions and key informant interviews with end users, community advisory boards, patient advocacy groups, and other stakeholders. The project team can then partner with an expert in app design, development, and programming. Last, the app should take into account the diversity of languages spoken in the local context and thereby allow children and caregivers to use the app in the language most suitable to them.

Given the paucity of robust evidence for using mobile apps to support adherence in children and adults, we suggest that future evaluations of mobile adherence apps to support children's adherence to TB treatment should include implementation science evaluations and randomized controlled trials. Further, such evaluations should measure both biomedical (eg, treatment success) and patient experience (eg, preferences and user satisfaction) outcomes.

Conclusion

Digital health technology is a burgeoning resource for supporting health systems and patients in the prevention, treatment, and care of a wide array of diseases. While such technologies hold promise for supporting the adherence of children diagnosed with TB disease in LMICs, where the burden is highest, no such technologies have been developed to date. In this paper, we presented some hypothetical scenarios of the daily struggles of caregivers and children who are tasked with adhering to TB treatment for prolonged periods of time. While we do not suggest that a mobile app will resolve all problems related to children's adherence to treatment, such a technology could compensate for and lessen some of the negative experiences children associate with taking TB treatment and perhaps even act as a buffer for health system shortcomings. It remains the task of the research community to develop such technologies and bring them to scale in the health systems in which they will have the greatest impact.

Authors' Contributions

RMM prepared a first draft of the manuscript, and all authors contributed substantially to the manuscript's conceptualization and subsequent revisions. We would like to thank authors for their interest, time, and support.

Conflicts of Interest

None declared.

References

1. Global tuberculosis report 2019. World Health Organization. Geneva, Switzerland: World Health Organization; 2019. URL: https://www.who.int/tb/publications/global_report/en/ [accessed 2020-10-21]
2. Handbook for the use of digital technologies to support tuberculosis medication adherence. World Health Organization. Geneva, Switzerland: World Health Organization; 2017. URL: https://www.who.int/tb/publications/2018/TB_medication_adherence_handbook_2018/en/ [accessed 2020-10-21]
3. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries--what has been shown to work? Glob Health Action 2014;7:25606 [FREE Full text] [doi: 10.3402/gha.v7.25606] [Medline: 25361730]
4. 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]
5. Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ. Mobile Apps for Increasing Treatment Adherence: Systematic Review. J Med Internet Res 2019 Jun 18;21(6):e12505 [FREE Full text] [doi: 10.2196/12505] [Medline: 31215517]
6. Vanleeuw L, Loveday M. Tuberculosis. In: Massyn N, Peer N, English R, Padarath A, Barron P, Day C, editors. District Health Barometer 2015/2016. Durban, South Africa: Health Systems Trust; 2016:181-208.
7. Guidance for National Tuberculosis Programmes on the Management of Tuberculosis in Children. 2nd ed. World Health Organization. Geneva, Switzerland: World Health Organization; 2014. URL: https://apps.who.int/iris/bitstream/handle/10665/112360/9789241548748_eng.pdf [accessed 2020-10-21]
8. Seddon JA, Schaaf HS. Drug-resistant tuberculosis and advances in the treatment of childhood tuberculosis. Pneumonia 2016;8:20 [FREE Full text] [doi: 10.1186/s41479-016-0019-5] [Medline: 28702299]
9. Goosby E. Out of the shadows: shining a light on children with tuberculosis. Int J Tuberc Lung Dis 2015 Dec;19(Suppl 1):1-2. [doi: 10.5588/ijtld.15.0411] [Medline: 26564533]
10. Zimri K, Casper R, Hoddinott G, Schaaf HS, Garcia-Prats AJ, Rose PC, et al. A novel approach for eliciting adolescent MDR-TB treatment tolerability: qualitative data from South Africa. Int J Tuberc Lung Dis 2020 Jan;24(1):43-47. [doi: 10.5588/ijtld.19.0207] [Medline: 32005306]
11. Meyerson K. Exploring caregivers' and health workers' perceptions on the effects of caregiver-child separation during long-term hospitalization for MDR-TB in the Western Cape: a qualitative study [thesis]. Stellenbosch University. 2019. URL: http://hdl.handle.net/10019.1/107320 [accessed 2020-10-21]
12. Wademan DT, Busakwe L, Nicholson TJ, van der Zalm M, Palmer M, Workman J, et al. Acceptability of a first-line anti-tuberculosis formulation for children: qualitative data from the SHINE trial. Int J Tuberc Lung Dis 2019 Dec;23(12):1263-1268 [FREE Full text] [doi: 10.5588/ijtld.19.0115] [Medline: 31931909]
13. Creary S, Chisolm D, Stanek J, Hankins J, O'Brien SH. A Multidimensional Electronic Hydroxyurea Adherence Intervention for Children With Sickle Cell Disease: Single-Arm Before-After Study. JMIR Mhealth Uhealth 2019 Aug;7(8):e13452 [FREE Full text] [doi: 10.2196/13452] [Medline: 31397291]
14. Ramsey RR, Holbein CE, Powers SW, Hershey AD, Kabbouche MA, O'Brien HL, et al. A pilot investigation of a mobile phone application and progressive reminder system to improve adherence to daily prevention treatment in adolescents and young adults with migraine. Cephalalgia 2018 Dec;38(14):2035-2044. [doi: 10.1177/0333102418756864] [Medline: 29635937]

15. Kosse RC, Bouvy ML, de Vries TW, Koster ES. Effect of a mHealth intervention on adherence in adolescents with asthma: A randomized controlled trial. Respir Med 2019 Mar;149:45-51. [doi: 10.1016/j.rmed.2019.02.009] [Medline: 30803885]
16. Whiteley L, Brown L, Mena L, Craker L, Arnold T. Enhancing health among youth living with HIV using an iPhone game. AIDS Care 2018;30(Suppl 4):21-33 [FREE Full text] [doi: 10.1080/09540121.2018.1503224] [Medline: 30626196]
17. Weisman O, Schonherz Y, Harel T, Efron M, Elazar M, Gothelf D. Testing the Efficacy of a Smartphone Application in Improving Medication Adherence, Among Children with ADHD. Isr J Psychiatry Relat Sci 2018;55(2):59-63 [FREE Full text] [Medline: 30368489]
18. Hightow-Weidman L, Muessig K, Knudtson K, Srivatsa M, Lawrence E, LeGrand S, et al. A Gamified Smartphone App to Support Engagement in Care and Medication Adherence for HIV-Positive Young Men Who Have Sex With Men (AllyQuest): Development and Pilot Study. JMIR Public Health Surveill 2018 Apr;4(2):e34 [FREE Full text] [doi: 10.2196/publichealth.8923] [Medline: 29712626]
19. LeGrand S, Muessig KE, Platt A, Soni K, Egger JR, Nwoko N, et al. Epic Allies, a Gamified Mobile Phone App to Improve Engagement in Care, Antiretroviral Uptake, and Adherence Among Young Men Who Have Sex With Men and Young Transgender Women Who Have Sex With Men: Protocol for a Randomized Controlled Trial. JMIR Res Protoc 2018 Apr;7(4):e94 [FREE Full text] [doi: 10.2196/resprot.8811] [Medline: 29622527]
20. DiStefano MJ, Schmidt H. mHealth for Tuberculosis Treatment Adherence: A Framework to Guide Ethical Planning, Implementation, and Evaluation. Glob Health Sci Pract 2016 Jun;4(2):211-221. [doi: 10.9745/GHSP-D-16-00018] [Medline: 27353615]
21. 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;74(Suppl 1):S37-S43 [FREE Full text] [doi: 10.1097/QAI.0000000000001202] [Medline: 27930610]
22. Lester R, Park JJ, Bolten LM, Enjetti A, Johnston JC, Schwartzman K, et al. Mobile phone short message service for adherence support and care of patients with tuberculosis infection: Evidence and opportunity. J Clin Tuberc Other Mycobact Dis 2019 Aug;16:100108 [FREE Full text] [doi: 10.1016/j.jctube.2019.100108] [Medline: 31720432]
23. Curtis K, Lebedev A, Aguirre E, Lobitz S. A Medication Adherence App for Children With Sickle Cell Disease: Qualitative Study. JMIR Mhealth Uhealth 2019 Jun;7(6):e8130 [FREE Full text] [doi: 10.2196/mhealth.8130] [Medline: 31215518]
24. Ahmed I, Ahmad NS, Ali S, Ali S, George A, Saleem DH, et al. Medication Adherence Apps: Review and Content Analysis. JMIR Mhealth Uhealth 2018 Mar;6(3):e62 [FREE Full text] [doi: 10.2196/mhealth.6432] [Medline: 29549075]
25. Abraham C, Sheeran P. The health belief model. In: Conner M, Norman P, editors. Predicting health behaviour: research and practice with social cognition models. 2nd ed. New York, NY: Open University Press; 2005:28-80.
26. Ajzen I. Models of human social behavior and their application to health psychology. Psychology and Health 1998 Jul;13(4):735-739 [FREE Full text] [doi: 10.1080/08870449808407426]
27. Skinner BF. Science and human behavior. New York, NY: Macmillan; 1953.
28. Carpenter CJ. A meta-analysis of the effectiveness of health belief model variables in predicting behavior. Health Commun 2010 Dec;25(8):661-669. [doi: 10.1080/10410236.2010.521906] [Medline: 21153982]
29. Morris JB, Tisi DA, Tan DCT, Worthington JH. Development and Palatability Assessment of Norvir® (Ritonavir) 100 mg Powder for Pediatric Population. Int J Mol Sci 2019 Apr 06;20(7):1718 [FREE Full text] [doi: 10.3390/ijms20071718] [Medline: 30959935]
30. Demonceau J, Ruppar T, Kristanto P, Hughes DA, Fargher E, Kardas P, et al. Identification and assessment of adherence-enhancing interventions in studies assessing medication adherence through electronically compiled drug dosing histories: a systematic literature review and meta-analysis. Drugs 2013 May;73(6):545-562. [doi: 10.1007/s40265-013-0041-3] [Medline: 23588595]
31. Naidoo P, Simbayi L, Labadarios D, Ntsepe Y, Bikitsha N, Khan G, et al. Predictors of knowledge about tuberculosis: results from SANHANES I, a national, cross-sectional household survey in South Africa. BMC Public Health 2016 Mar;16:276 [FREE Full text] [doi: 10.1186/s12889-016-2951-y] [Medline: 26987759]
32. Cramm JM, Finkenflügel HJM, Møller V, Nieboer AP. TB treatment initiation and adherence in a South African community influenced more by perceptions than by knowledge of tuberculosis. BMC Public Health 2010 Feb;10:72 [FREE Full text] [doi: 10.1186/1471-2458-10-72] [Medline: 20163702]
33. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. Pediatrics 2008 Aug;122(2):e305-e317. [doi: 10.1542/peds.2007-3134] [Medline: 18676516]
34. Brown SJ, Lieberman DA, Germeny BA, Fan YC, Wilson DM, Pasta DJ. Educational video game for juvenile diabetes: results of a controlled trial. Med Inform (Lond) 1997;22(1):77-89. [Medline: 9183781]
35. Story A, Aldridge RW, Smith CM, Garber E, Hall J, Ferenando G, et al. Smartphone-enabled video-observed versus directly observed treatment for tuberculosis: a multicentre, analyst-blinded, randomised, controlled superiority trial. Lancet 2019 Mar;393(10177):1216-1224 [FREE Full text] [doi: 10.1016/S0140-6736(18)32993-3] [Medline: 30799062]
36. Nsengiyumva NP, Mappin-Kasirer B, Oxlade O, Bastos M, Trajman A, Falzon D, et al. Evaluating the potential costs and impact of digital health technologies for tuberculosis treatment support. Eur Respir J 2018 Nov;52(5):1801363. [doi: 10.1183/13993003.01363-2018] [Medline: 30166325]

Abbreviations

DR: drug-resistant
DS: drug-susceptible
FAQ: frequently asked question
LMIC: low- and middle-income country
SCD: sickle cell disease
TB: tuberculosis
WHO: World Health Organization

Edited by G Eysenbach; submitted 06.04.20; peer-reviewed by S Browne, J Stanek; comments to author 12.06.20; revised version received 03.09.20; accepted 07.09.20; published 11.11.20.

Please cite as:
Morse RM, Myburgh H, Reubi D, Archey AE, Busakwe L, Garcia-Prats AJ, Hesseling AC, Jacobs S, Mbaba S, Meyerson K, Seddon JA, van der Zalm MM, Wademan DT, Hoddinott G
Opportunities for Mobile App–Based Adherence Support for Children With Tuberculosis in South Africa
JMIR Mhealth Uhealth 2020;8(11):e19154
URL: https://mhealth.jmir.org/2020/11/e19154
doi:10.2196/19154
PMID:33174850

Viewpoint

Evaluating Patient-Centered Mobile Health Technologies: Definitions, Methodologies, and Outcomes

Courtenay Bruce[1], JD, MA; Patricia Harrison[1], BSN, RN, MBA; Charlie Giammattei[2], BA; Shetal-Nicholas Desai[3,4], DPM, MBA; Joshua R Sol[3,4], BSc; Stephen Jones[5,6], MD, MSHI; Roberta Schwartz[3], PhD

[1]System Quality & Patient Safety, Houston Methodist System, Houston, TX, United States

[2]CareSense, MedTrak, Inc, Conshohocken, PA, United States

[3]Center for Innovation, Houston Methodist Hospital, Houston, TX, United States

[4]Information Technology Division, Houston Methodist Hospital, Houston, TX, United States

[5]Center for Outcomes Research, Houston Methodist Research Institute, Houston, TX, United States

[6]Department of Surgery, Houston Methodist Hospital, Houston, TX, United States

Corresponding Author:
Courtenay Bruce, JD, MA
System Quality & Patient Safety
Houston Methodist System
6565 Fannin Street
Houston, TX, 77030
United States
Phone: 1 2816209040
Email: crbruce@houstonmethodist.org

Abstract

Several recently published studies and consensus statements have demonstrated that there is only modest (and in many cases, low-quality) evidence that mobile health (mHealth) can improve patient clinical outcomes such as the length of stay or reduction of readmissions. There is also uncertainty as to whether mHealth can improve patient-centered outcomes such as patient engagement or patient satisfaction. One principal challenge behind the "effectiveness" research in this field is a lack of common understanding about what it means to be effective in the digital space (ie, what should constitute a relevant outcome and how best to measure it). In this viewpoint, we call for interdisciplinary, conceptual clarity on the definitions, methodologies, and patient-centered outcomes frequently used in mHealth research. To formulate our recommendations, we used a snowballing approach to identify relevant definitions, outcomes, and methodologies related to mHealth. To begin, we drew heavily upon previously published detailed frameworks that enumerate definitions and measurements of engagement. We built upon these frameworks by extracting other relevant measures of patient-centered care, such as patient satisfaction, patient experience, and patient activation. We describe several definitional inconsistencies for key constructs in the mHealth literature. In an effort to achieve clarity, we tease apart several patient-centered care outcomes, and outline methodologies appropriate to measure each of these patient-care outcomes. By creating a common pathway linking definitions with outcomes and methodologies, we provide a possible interdisciplinary approach to evaluating mHealth technologies. With the broader goal of creating an interdisciplinary approach, we also provide several recommendations that we believe can advance mHealth research and implementation.

(JMIR Mhealth Uhealth 2020;8(11):e17577) doi:10.2196/17577

KEYWORDS

innovation; health care; digital technology; digital interventions; patient-facing technologies; patient-centered care; patient centeredness; patient experience; patient engagement; patient activation; quality; effectiveness; quality improvement; information technologies; outcomes; readmissions; length of stay; patient adherence

Background

Mobile health (mHealth) is defined by the Word Health Organization as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistance, and other wireless devices" [1,2]. mHealth is considered the future of health care [3,4], and many health care organizations have embraced mHealth as part of their patient-centered initiatives. Specifically, according to a US News & World report, 18 of the top 20 medical centers

have adopted and widely celebrated mHealth technologies, at least in a review we conducted of their websites [5]. The National Institutes of Health funding for developing, testing, and implementing mHealth interventions grew from US $16.8 million in 2014 to US $39.4 million in 2018 [6].

Despite the growth and hype of mHealth technologies, there is a paucity of effectiveness evidence in the literature to support such widespread implementation [6]. Several recently published studies and consensus statements have demonstrated that there is only modest (and in many cases, low-quality) evidence that mHealth can improve patient clinical outcomes such as lengths of stay or reduction of readmissions [6-9]. There is also uncertainty as to whether mHealth can improve patient-centered outcomes such as patient engagement or patient satisfaction [6-10]. One principal challenge behind the "effectiveness" research is a lack of shared understanding about what it means to be effective in the digital space (ie, what should constitute a relevant outcome and how best to measure it) [2]. This lack of a shared understanding can likely be attributed to the multidisciplinary, multifaceted nature of mHealth, often involving disciplines such as engineering, data science, systems science, human-computer interaction, behavioral sciences, public health, and medicine [10-13]. Each discipline tends to frame its effectiveness research in terms of its own specialized knowledge [13], thereby limiting the generalizability of research results across domains.

In this viewpoint, we call for interdisciplinary, conceptual clarity on the definitions, methodologies, and patient-centered outcomes frequently used in mHealth research. There have been recent, important calls in the literature for conceptual clarity about engagement [6,10-13]; however, as we demonstrate below, engagement is only one facet of patient-centered care [14] and it is not always clear that sustained engagement is required to achieve mHealth outcomes [2]. Stated differently, a shared understanding about engagement is necessary but not sufficient for an interdisciplinary research approach. Conceptual clarity on engagement can only go so far to reduce the current fragmentation of research efforts [12]. Instead, what is needed is consideration beyond one measure of patient-centered care to include several measures of patient-centered care and the methods for evaluating them. By creating a common pathway linking definitions with outcomes and methodologies, we hope to draw a comprehensive (but not necessarily exhaustive) outline of a possible interdisciplinary approach to evaluating mHealth technologies.

To formulate our recommendations, we used a snowballing approach to identify relevant definitions, outcomes, and methodologies related to mHealth. To begin, we drew heavily upon the work of Perski et al [12], Yardley et al [6,13], and Short et al [10], among others [14,15], whose detailed frameworks enumerate the definitions and measurements of engagement [15]. We built upon their frameworks by extracting other relevant measures of patient-centered care, such as patient satisfaction, patient experience, and patient activation, drawing on the quality and patient safety literature—a literature base that has, thus far, to our knowledge, been largely untapped in the mHealth context.

Quality health care refers to care that is safe, effective, patient-centered, timely, efficient, and equitable [11,16]. In the mHealth context, most of the work comes from informaticians rather than health care quality practitioners. Theory-based informatics research must be informed by the frontline real-world, hospital environment where health care quality takes place, because patients are different than consumers or research participants [11,17]. Patients experience a constellation of complex, emotionally laden perspectives during their use of mHealth technology that may not be considered when conducting informatics research outside of the hospital setting [18]. The field of health care quality can help advance mHealth research in its evidence-based emphasis on the patient, his or her experiences, and how the continuum of care can influence outcomes that matter to the patient [19,20].

A data extraction table was used to sort, explore, and synthesize existing research (see Multimedia Appendix 1). We filtered and coded our findings based on whether a definition was proposed, whether outcome measures were discussed (and what the outcome measures were), and methodologies used to assess the outcomes. We reran the search queries that were performed in 5 recent, frequently cited systematic reviews [6,12,14,21,22] using the Medline, PsycINFO, PubMed, and Google Scholar databases. Although we did not intend to perform a systematic review for this Viewpoint, we believe that this research approach was comprehensive and could be used to formulate recommendations based on gaps and inconsistencies in the literature.

Defining Patient-Centered mHealth Technologies

A survey of the literature highlights the myriad terms used in mHealth and the various ways in which they are defined. The terms "mHealth," "telehealth," "eHealth," and "digital technologies" are often treated synonymously [23], although, in reality, they could be different in that eHealth or digital technologies can encompass devices that are not supported by mobile means, such as hospital check-in or registration portals to more advanced technologies designed to enhance patient understanding through education and communication such as Smart Boards and Smart TVs located on hospital units [24].

Digital behavior change interventions (DBCIs) are a subset of eHealth, defined as "a product or service that uses computer technology to promote behavior change," which can be delivered through computer programs, websites, mobile phones as text message, smartphone apps, or wearable devices [6].

Patient-centered care is a health care quality indicator proposed by the Institute of Medicine (IOM) [25]. Patient-centered care has generally been poorly defined, with authors often conflating several distinct concepts such as using the term "patient-centered care" when they are usually referring to a specific outcome measure of patient-centered care [26]. As the IOM explains, patient-centered care is care that is "respectful of and responsive to individual patient preferences, needs, and values, and ensuring that patient values guide all clinical decisions" [25]. Through this lens, patient-centered care can be conceptualized as an

overarching, broad concept that puts emphasis on respecting patients as a means of honoring their dignity and worth [26,27]. By extension, patient-centered mHealth technologies can be defined as technologies supported by mobile devices designed to promote patient-centered care. There are several outcome measures for patient-centered care.

Patient-Centered Care Outcomes for Patient-Centered mHealth Technologies

"Patient experience" is likely the strongest, most proximate outcome measure of patient-centered care. Patient experience refers to any process observable by patients, including subjective experiences (eg, pain control) and their objective experiences (eg, wait time) [28,29]. A primary feature of patient experience is that it reflects actual health care experiences [30]. Another important feature of patient experience is that it refers to a time sequence. Specifically, patient experience can refer to the first touchpoint within an episode of care (eg, being assessed for a left knee replacement) to the last touchpoint within that episode (eg, last follow-up appointment following surgery for a left knee replacement). Alternatively, patient experience can refer to the whole continuum of care—their first encounter with Hospital A to their last encounter with Hospital A.

"Patient engagement" and "patient activation" are two outcome measures of patient-centered care that are distinct from patient experience. In particular, "patient engagement" has been defined in various ways in the literature [14,30,31]. Recent proposals for integrative definitions of engagement have defined engagement as consisting of objective and subjective components [6,12,13]. The objective component is the extent (eg, amount, frequency, duration, and depth) of usage of the mHealth technology [12]. With respect to what constitutes sufficient objective engagement, the literature demonstrates a palpable lack of consensus [15,31]. Some researchers have proposed that a certain empirical threshold of engagement must be met to show sufficient engagement with the intervention to achieve intended outcomes [13], whereas others suggest that one critical point of engagement or fluid, ebbing, and flowing engagement may be sufficient [32,33]. The subjective element of engagement is often characterized by the user's attention, interest, and affect—their overall experience in engaging with the specific mHealth technology [10,12].

"Patient activation" is a patient-centered care outcome measure that refers to patients' "willingness and ability to take independent actions to manage their health," such as avoiding health-damaging behaviors and adopting healthy lifestyle choices, including exercising regularly, eating well, or monitoring their glucose levels [34]. DBCIs usually attempt to evaluate patient activation [34-36], although they may indicate that they are designed to assess adherence or engagement. "Adherence" is often confused with "engagement" and has been defined in at least three separate ways: adherence could refer to whether the intervention is used as intended by the developers [10,32], the usage of mHealth [37] (which is more accurately called engagement), or the patient's willingness and ability to adhere to the recommendations provided by their physician or health care provider (which is likely the most common use of the word "adherence" in medico-legal parlance) [35,38-40].

Taken together, patient engagement typically refers to patients' engagement with an intervention itself, whereas patient activation refers to patients' physical and mental health-related activities based on what they learned from an intervention [34,41]. Patients can be engaged by, for example, reading digital messages on the importance of exercising and still not be activated to start exercising [34].

Patient engagement and activation are distinct outcome measures from patient experience in that patient experience refers to patients' perceptions of others' actions, whereas patient engagement and activation refer to patients' actions [27,42]. Patient experience, engagement, and activation are all measures of patient-centered care in that they put emphasis on the patient playing an integral role in their outcomes [43,44], with the ultimate decision-making authority resting with patients. Further, all three concepts can be evaluated empirically.

Finally, "patient satisfaction" is the most attenuated outcome measure of patient-centered care. Patient satisfaction is a term that is often used in health care quality parlance and yet is frequently misunderstood [27] owing to one key feature of satisfaction that is often ignored: satisfaction has little to do with quality [22,27,29]. Patient satisfaction only refers to whether patients' expectations were met [22,27]. Patients can be satisfied with care that is low quality and yet be dissatisfied with high-quality care. For this reason, it is incorrect to say that the Hospital Consumer Assessment of Healthcare Provider and Systems survey, the first standardized, publicly reported survey of patient perspectives [45], measures patient satisfaction when, in actuality, the survey uses patient experience as its primary measure for patient-centered care [27,28,46].

As the above discussion illustrates, to date, we have not achieved a shared understanding of important mHealth constructs, or how to conceptualize and operationalize them [10]. Thus, if the mHealth community is to continue to promote the use of patient-centered mHealth technologies (ie, mHealth technologies that are designed with the goal of promoting patient-centered care, as measured by patient experience, patient engagement, patient activation, or patient satisfaction), then we need precision in our terminology to extrapolate generalizable, transferable results [6,31].

Methodologies Used to Empirically Assess Patient-Centered Outcomes

To validly measure a concept, there must be a tight linkage between the patient-centered care construct (ie, the outcome measure) and items developed for measurement [47]. Below, we describe several methodologies that can assess each patient-centered outcome.

Patient Experience

There are several qualitative methodologies that are ideal for assessing patient experience. Focus groups, semistructured interviews, observational studies, patient journey mapping, and walk-throughs are all appropriate to assess patient experience

[21,29,48,49]. Semistructured interviews (where the interviewer uses a structured guide to aid conversation) are most helpful to elicit patients' experiences, as well as to elicit particular informational and decisional needs. For example, consider an mHealth technology that provides educational messages (through text and email) to prepare and inform patients about their upcoming surgery. Asking patients to reflect on whether they had gaps in knowledge after completion of the module (eg, gaps in knowledge about surgery lifestyle impacts, surgical risks/benefits, and technical aspects of the surgery, or recovery trajectories and activities) would help elicit patients' informational needs [29].

To assess patient experience, researchers should aim to elicit both definitional components of patient experience: (a) patients' subjective and objective assessments of what occurred, and (b) patients' observations across the full sequence of time [28,45]. Thus, if only one touchpoint of care within an episode is evaluated—such as patient experiences using a digital check-in registration system during a scheduled surgery—then it likely cannot be said that patient experience was fully assessed. A more precise methodology would be one in which teams systematically evaluate patient experiences through all touchpoints using walk-throughs or patient journey mapping, starting from appointment scheduling, to the registration check-in, to the digital navigation system that shows patients how to get to a particular department, to the Smart TVs or Smart Boards within the patients' hospital rooms, and all the way to the patients' beside tablet-based and electronic health record solutions [50-52]. The goal of walk-throughs and journey mapping is to ask patients what they are feeling, seeing, and experiencing as they move from, say, the registration portal to a patient room [50].

Patient Engagement

Short and colleagues [10] describe several methodologies that are appropriate to evaluate patient engagement. To assess patient engagement, researchers should aim to elicit both definitional components of patient engagement: the objective component, as the extent (eg, amount, frequency, duration, and depth) of usage of the mHealth technology, as well users' subjective assessment in using the mHealth technology [13]. The objective component is likely best evaluated using quantitative measures such as the number of login attempts, the time spent on a technology, the time spent reading a particular message or conducting an e-module, or the amount of bidirectional communication between a patient and provider using an mHealth technology [11,13,53-58]. The subjective component of engagement can likely best be evaluated using qualitative methodologies such as semistructured interviews and focus groups, which tend to be most appropriate for teasing out themes, as well as users' beliefs, narratives, and perceptions [23].

Patient Activation

There is a heavy behavioral dimension to patient activation—what the patient does in response to the intervention in terms of his or her health-related activities—which can be assessed quantitatively [58,59]. Several studies that claim to measure patient engagement are arguably instead measuring patient activation [60-62]. Whether a patient had higher medication adherence [63,64], higher levels of physical exercise [9,65,66], or improved diabetes management as a result of the mHealth intervention [67,68] can be considered patient activation if the patient took healthy actions based on what he/she learned through the mHealth technology. These medical, clinically based health outcomes can be empirically derived using electronic medical records. There are also validated, reliable quantitative measures available to evaluate patient activation [58,59].

Patient Satisfaction

Patient satisfaction is generally considered to be the most attenuated outcome measure for patient-centered mHealth technologies and thus can likely be assessed from a purely descriptive [69] quantitative point of view, devoid of any thematic nuancing that qualitative measures can afford. There are numerous quantitative measures available to evaluate patient satisfaction, although the validity and reliability of the instruments have been a point of debate [69-80].

Table 1 outlines a common pathway linking definitions with outcomes and methodologies. The table is not intended to be exhaustive but is instead designed to provide a robust set of patient-centered constructs, outcome measures, and methodologies. In this summary, we refrained from including strictly objective, physiological measures because the patient-centered constructs depend heavily on the subjective experience of patients, which physiological measures often cannot elicit.

Table 1. Patient-centered mobile health (mHealth) technologies: outcome measures, methodologies, and definitions.

Outcome measures	Definitions	Methodologies		Example mHealth Technologies
		Qualitative	Quantitative	
Patient experience	Any process observable by patients, including subjective experiences (eg, pain control) and their objective experiences (eg, wait time). Must refer to the entire sequence in the care episode or full continuum of care—from the first to the last touchpoint.	Semistructured interviews Think aloud exercises Focus groups	Walk-throughs (eg, the Walk Through Tool [48]) Patient Journey Mapping Tool [49,50] Impact of Assistive Devices Scale (PIADS) [51]	Touchscreen kiosks for registration or check-in Digital navigation systems to help patients navigate the hospital Smart TVs in hospital rooms
Patient engagement	The extent (eg, amount, frequency, duration, and depth) of usage of the mHealth technology, coupled with the user's subjective assessment in using the mHealth technology.	Semistructured interviews Think aloud exercises Focus groups	Self-report questionnaires (eg, the eHealth Engagement Scale and the Digital Behavior Change Intervention Engagement Scale [52]) Usability and acceptability scales (eg, The mHealth App Usability Questionnaire [MAUQ] [53]) Social Networking Time Use Scale (SONTUS) [54] Facebook Intensity Scale (FBI Scale) [55] Media and Technology Usage and Attitudes Scale (MTUAS) [56] Chinese Internet Gaming Disorder Scale [57] Number of logins; time spent on mHealth; time spent reading a message; number of monitoring questions for which there was a response	Two-way bidirectional communication on a communication or education mobile platform Secure SMS text message/email/push notifications, self-scheduling, medical record access, patient-provider messaging, and bill pay
Patient activation	Willingness and ability to take independent actions to manage their health	Focus groups, semistructured interviews	Patient Activation Measure (PAM) [58,59]	Digital behavior change interventions Patient biosensor monitoring devices (eg, glucose monitoring kits)
Patient satisfaction	Whether a patient's expectations were met.	Semistructured interviews	Short Assessment of Patient Satisfaction (SAPS) [60] Risser Patient Satisfaction Scale [61] Patient Satisfaction Questionnaire (PSQ) (multiple iterations) [62] Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST) [63]	Patient-reported outcomes collection via mHealth

Recommendations

Above, we have described several definitional inconsistencies for key constructs in the mHealth literature. In an effort to achieve clarity, we teased apart several patient-centered care outcomes, and we outline methodologies appropriate to measure each of the patient-care outcomes (Table 1). In what follows, we provide recommendations for evaluating mHealth patient-centered technologies. Our recommendations relate to the patient-centered constructs in that we advocate for patients taking center stage in mHealth research and collaboration efforts to enhance patient experience, patient engagement, patient activation, and patient satisfaction.

Recommendation 1: Use Patients When Assessing Patient-Centered Care mHealth Technologies

Patient participation is a central tenet of ethically driven research and product development [81-83]. Unfortunately, much of mHealth research only engages patients in the beginning phases during agenda-setting and protocol development [83]. Systematic reviews have found that little research involves patients throughout the development, implementation, and modification phases of mHealth patient-centered technologies [83,84]. Reviews have also found that patient participation is often treated as a tokenistic measure, one in which patients' feedback is used primarily as a means of "rubberstamping" to secure funding or to approve a previously chosen decision made

by the research or design team, rather than using patients as principal drivers of decision-making [83,84].

The main reason that patient participation is integral to conducting ethically driven mHealth research and product development is because patients are unique [42,85]; they should not be viewed as the same type of customers found in other sectors [6]. For instance, customers of any other good or service have luxuries that patients may not have. Patients become consumers, arguably not by choice but rather by need. Patients' preferences are largely unknown when the good or service is used, and they have limited channels of communication and limited control [42]. Patients likely experience physical or emotional impairments such as fears, grief, and anxiety while using the service [86]. All of these factors likely suggest that a particular patient's informational and decision-making needs are different from those that they would have outside of a health care context [86-89]. Patient participation and active engagement are the classic tenets of task, user, representation, and functional analysis used to inform system designers and leadership that are contemplating implementing solutions to patient-centered problems [87-89].

Recommendation 2: Create an International Collaboration to Enhance the Quality and Effectiveness of mHealth Technologies

mHealth technologies are the future for patient-centered care [3,74]. However, the extent to which mHealth technologies influence or impact patient-centered care is unclear [6], which is likely owing to the imprecision in definitions, methodological approaches, and disjointed interests of multiple stakeholders within health care organizations and industry [2].

We contend that some form of self-regulation or an internationally used assessment framework is needed to ensure that quality standards are met before wide-scale dissemination of any patient-centered mHealth technology [2]. The US Food and Drug Administration has taken a passive approach, explicitly applying its regulatory oversight to only those software functions that are medical devices and whose functionality could pose a risk to a patient safety if the device were to not function as intended [90]. The European Commission mHealth Green Paper does not give any recommendations [91]. The Health Technology Assessment Agency and its collaborative networks have discussed the importance of a collaborative approach, but, at the time of this writing, have failed to provide some form of a comprehensive evaluative framework [92,93].

When patient decision aids were being developed and implemented at a rapid pace without high-quality evidence, an international collaboration among researchers, practitioners, and stakeholders was instituted to enhance the quality and effectiveness of patient decision aids, called the International Patient Decision Aids (IPDAS) Collaboration [94]. The group established (and routinely revises) an evidence-informed framework, which outlines actions that patient decision-aid developers should take in terms of content-writing, development, implementation, and evaluation. The IPDAS Collaboration "grades" decision aids based on whether and to what degree the aids meet parameters, and the score can be used in marketing and evaluating the patient decision aid [94].

An international collaboration similar to that used for patient decision aids is, we believe, appropriate for the regulation and systematic evaluation of patient-centered mHealth technologies. Similar to patient decision aids, the successful implementation of patient-centered mHealth requires multidisciplinary teams from academia, industry, and health care management sectors, along with patients and consumers working collaboratively to maintain the requisite medical, statistical, information technology, patient-centered, and research expertise necessary to implement and evaluate mHealth technologies [95], which an international collaboration would afford. There are several actions a large collaboration could take to encourage high-quality development and dissemination of digital technologies.

First, because of its large scale, an international collaboration would be well-positioned to help address barriers to electronic interoperability issues that stem from disparate proprietary digital health record systems by, for example, creating digital health data exchange platforms to standardize data [95]. Second, a collaboration could disseminate practical advice on how organizations can use their foundational digital systems to leverage existing capabilities for achieving coordination through bolt-on, incremental development of digital technologies. Third, an international collaboration could build upon our work to develop exhaustive criteria and methodology standards for how to design, produce, implement, and evaluate digital technologies.

Conclusion

In this Viewpoint, we called for interdisciplinary, conceptual clarity on the definitions, methodologies, and patient-centered outcomes frequently used in mHealth research. In doing so, we advocate for consideration of several measures of patient-centered care, and we outline various methods for evaluating them. By creating a common pathway linking definitions with outcomes and methodologies, we provide a possible interdisciplinary approach to evaluating mHealth technologies.

To that end of creating an interdisciplinary approach, we also provide several recommendations that we believe can advance mHealth research and implementation. For instance, if an international collaboration were created to develop evaluative criteria, using the guidance provided here to ground criteria development, then low-quality digital technologies would likely be excluded. Transparency and precision would be promoted, large-scale published evidence would be encouraged [3], and mHealth technologies could finally flourish within a high-quality, patient-centered landscape.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data Abstraction Table.
[XLSX File (Microsoft Excel File), 37 KB - mhealth_v8i11e17577_app1.xlsx]

References

1. Frequently asked questions on Global Task Force on digital health for TB and its work. World Health Organization. URL: http://www.who.int/tb/areas-of-work/digital-health/faq/en/ [accessed 2020-11-03]
2. 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]
3. Smith B, Sverdlov A. Digital Technology: The Future Is Bright. Clin Pharmacol Ther 2018 Jul;104(1):9-11. [doi: 10.1002/cpt.1092] [Medline: 29890004]
4. Alami H, Gagnon M, Fortin J. Digital health and the challenge of health systems transformation. Mhealth 2017;3:31. [doi: 10.21037/mhealth.2017.07.02] [Medline: 28894741]
5. Health News Articles. US News & World Report Health. URL: https://health.usnews.com/ [accessed 2020-11-03]
6. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and Evaluating Digital Interventions to Promote Behavior Change in Health and Health Care: Recommendations Resulting From an International Workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
7. Donevant SB, Estrada RD, Culley JM, Habing B, Adams SA. Exploring app features with outcomes in mHealth studies involving chronic respiratory diseases, diabetes, and hypertension: a targeted exploration of the literature. J Am Med Inform Assoc 2018 Oct 01;25(10):1407-1418 [FREE Full text] [doi: 10.1093/jamia/ocy104] [Medline: 30137383]
8. Aitken M, Lyle J. Patient Adoption of mHealth: Use, Evidence and Remaining Barriers to Mainstream Acceptance. Parsippany, NJ: IMS Institute for Healthcare Informatics; 2015.
9. Patnode CD, Evans CV, Senger CA, Redmond N, Lin JS. Behavioral Counseling to Promote a Healthful Diet and Physical Activity for Cardiovascular Disease Prevention in Adults Without Known Cardiovascular Disease Risk Factors: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA 2017 Jul 11;318(2):175-193. [doi: 10.1001/jama.2017.3303] [Medline: 28697259]
10. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring Engagement in eHealth and mHealth Behavior Change Interventions: Viewpoint of Methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
11. Shekelle PG, Pronovost PJ, Wachter RM, Taylor SL, Dy SM, Foy R, et al. Advancing the science of patient safety. Ann Intern Med 2011 May 17;154(10):693-696. [doi: 10.7326/0003-4819-154-10-201105170-00011] [Medline: 21576538]
12. 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]
13. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
14. Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF. The Impact of Information Technology on Patient Engagement and Health Behavior Change: A Systematic Review of the Literature. JMIR Med Inform 2016 Jan 21;4(1):e1 [FREE Full text] [doi: 10.2196/medinform.4514] [Medline: 26795082]
15. 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]
16. About health care quality. US Department of Health and Human Services Office of Disease Prevention and Health Promotion. URL: https://health.gov/our-work/health-care-quality/about-health-care-quality [accessed 2020-11-03]
17. Davidoff F, Dixon-Woods M, Leviton L, Michie S. Demystifying theory and its use in improvement. BMJ Qual Saf 2015 Mar;24(3):228-238 [FREE Full text] [doi: 10.1136/bmjqs-2014-003627] [Medline: 25616279]
18. Batalden M, Batalden P, Margolis P, Seid M, Armstrong G, Opipari-Arrigan L, et al. Coproduction of healthcare service. BMJ Qual Saf 2016 Jul;25(7):509-517 [FREE Full text] [doi: 10.1136/bmjqs-2015-004315] [Medline: 26376674]
19. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. Health Aff (Millwood) 2008;27(3):759-769. [doi: 10.1377/hlthaff.27.3.759] [Medline: 18474969]
20. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12(6):573-576 [FREE Full text] [doi: 10.1370/afm.1713] [Medline: 25384822]
21. Newell S, Jordan Z. The patient experience of patient-centered communication with nurses in the hospital setting: a qualitative systematic review protocol. JBI Database System Rev Implement Rep 2015 Jan;13(1):76-87. [doi: 10.11124/jbisrir-2015-1072] [Medline: 26447009]
22. Batbaatar E, Dorjdagva J, Luvsannyam A, Savino MM, Amenta P. Determinants of patient satisfaction: a systematic review. Perspect Public Health 2017 Mar;137(2):89-101. [doi: 10.1177/1757913916634136] [Medline: 27004489]

23. Njoroge M, Zurovac D, Ogara EAA, Chuma J, Kirigia D. Assessing the feasibility of eHealth and mHealth: a systematic review and analysis of initiatives implemented in Kenya. BMC Res Notes 2017 Feb 10;10(1):90 [FREE Full text] [doi: 10.1186/s13104-017-2416-0] [Medline: 28183341]
24. Lee S, Cho Y, Kim S. Mapping mHealth (mobile health) and mobile penetrations in sub-Saharan Africa for strategic regional collaboration in mHealth scale-up: an application of exploratory spatial data analysis. Global Health 2017 Aug 22;13(1):63 [FREE Full text] [doi: 10.1186/s12992-017-0286-9] [Medline: 28830540]
25. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington DC: National Academies Press; 2001.
26. Tzelepis F, Sanson-Fisher RW, Zucca AC, Fradgley EA. Measuring the quality of patient-centered care: why patient-reported measures are critical to reliable assessment. Patient Prefer Adherence 2015;9:831-835. [doi: 10.2147/PPA.S81975] [Medline: 26150703]
27. Cleary PD. Evolving Concepts of Patient-Centered Care and the Assessment of Patient Care Experiences: Optimism and Opposition. J Health Polit Policy Law 2016 Aug;41(4):675-696. [doi: 10.1215/03616878-3620881] [Medline: 27127265]
28. Browne K, Roseman D, Shaller D, Edgman-Levitan S. Health Aff (Millwood) 2010 May;29(5):921-925. [doi: 10.1377/hlthaff.2010.0238] [Medline: 20439881]
29. Anhang Price R, Elliott MN, Zaslavsky AM, Hays RD, Lehrman WG, Rybowski L, et al. Examining the role of patient experience surveys in measuring health care quality. Med Care Res Rev 2014 Oct;71(5):522-554 [FREE Full text] [doi: 10.1177/1077558714541480] [Medline: 25027409]
30. Evans D, Hopewell-Kelly N, Kok M, White J. Synthesising conceptual frameworks for patient and public involvement in research - a critical appraisal of a meta-narrative review. BMC Med Res Methodol 2018 Oct 25;18(1):116 [FREE Full text] [doi: 10.1186/s12874-018-0572-0] [Medline: 30359226]
31. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inf Sci 2008 Apr;59(6):938-955. [doi: 10.1002/asi.20801]
32. Oliver S, Clarke-Jones L, Rees R, Milne R, Buchanan P, Gabbay J, et al. Involving consumers in research and development agenda setting for the NHS: developing an evidence-based approach. Health Technol Assess 2004 Apr;8(15):1-148, III [FREE Full text] [doi: 10.3310/hta8150] [Medline: 15080866]
33. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating Digital Health Interventions: Key Questions and Approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
34. 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]
35. Kelders SM, Van Gemert-Pijnen JEWC, Werkman A, Nijland N, Seydel ER. Effectiveness of a Web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. J Med Internet Res 2011 Apr 14;13(2):e32 [FREE Full text] [doi: 10.2196/jmir.1624] [Medline: 21493191]
36. Shippee ND, Domecq Garces JP, Prutsky Lopez GJ, Wang Z, Elraiyah TA, Nabhan M, et al. Patient and service user engagement in research: a systematic review and synthesized framework. Health Expect 2015 Oct;18(5):1151-1166 [FREE Full text] [doi: 10.1111/hex.12090] [Medline: 23731468]
37. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the Concept of Adherence to eHealth Technology: Systematic Review on When Usage Becomes Adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
38. Haywood C, Lanzkron S, Bediako S, Strouse JJ, Haythornthwaite J, Carroll CP, IMPORT Investigators. Perceived discrimination, patient trust, and adherence to medical recommendations among persons with sickle cell disease. J Gen Intern Med 2014 Dec;29(12):1657-1662 [FREE Full text] [doi: 10.1007/s11606-014-2986-7] [Medline: 25205621]
39. Martin LR, Williams SL, Haskard KB, Dimatteo MR. The challenge of patient adherence. Ther Clin Risk Manag 2005 Sep;1(3):189-199 [FREE Full text] [Medline: 18360559]
40. Tilson HH. Adherence or Compliance? Changes in Terminology. Ann Pharmacother 2004 Jan;38(1):161-162. [doi: 10.1345/aph.1d207]
41. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. Health Aff (Millwood) 2013 Feb;32(2):207-214. [doi: 10.1377/hlthaff.2012.1061] [Medline: 23381511]
42. Boissy A. Better communication makes better physicians. NEJM Catalyst. URL: https://catalyst.nejm.org/doi/full/10.1056/CAT.17.0377 [accessed 2020-11-03]
43. Dykes PC, Rozenblum R, Dalal A, Massaro A, Chang F, Clements M, et al. Prospective Evaluation of a Multifaceted Intervention to Improve Outcomes in Intensive Care: The Promoting Respect and Ongoing Safety Through Patient Engagement Communication and Technology Study. Crit Care Med 2017 Aug;45(8):e806-e813. [doi: 10.1097/CCM.0000000000002449] [Medline: 28471886]
44. Black N, Varaganum M, Hutchings A. Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery. BMJ Qual Saf 2014 Jul;23(7):534-542. [doi: 10.1136/bmjqs-2013-002707] [Medline: 24508681]

45. HCAHPS: Patients' Perspectives of Care Survey. Centers for Medicare & Medicaid Services. 2017 Dec 21. URL: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalHCAHPS.html [accessed 2020-11-03]
46. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. BMJ Open 2013 Jan 03;3(1):e001570 [FREE Full text] [doi: 10.1136/bmjopen-2012-001570] [Medline: 23293244]
47. Hamric AB. Empirical research on moral distress: issues, challenges, and opportunities. HEC Forum 2012 Mar 4;24(1):39-49. [doi: 10.1007/s10730-012-9177-x] [Medline: 22476738]
48. Gustafson D. Walk-through Tool. Institute for Healthcare Improvement. URL: http://www.ihi.org/resources/Pages/Tools/Walkthrough.aspx [accessed 2020-11-03]
49. Trebble TM, Hansi N, Hydes T, Smith MA, Baker M. Process mapping the patient journey: an introduction. BMJ 2010 Aug 13;341:c4078. [doi: 10.1136/bmj.c4078] [Medline: 20709715]
50. Antonacci G, Reed JE, Lennox L, Barlow J. The use of process mapping in healthcare quality improvement projects. Health Serv Manage Res 2018 May;31(2):74-84. [doi: 10.1177/0951484818770411] [Medline: 29707978]
51. Jutai J, Day H. Psychosocial Impact of Assistive Devices Scale (PIADS). Technol Disab 2002 Sep 29;14(3):107-111. [doi: 10.3233/tad-2002-14305]
52. Perski O, Lumsden J, Garnett C, Blandford A, West R, Michie S. Assessing the Psychometric Properties of the Digital Behavior Change Intervention Engagement Scale in Users of an App for Reducing Alcohol Consumption: Evaluation Study. J Med Internet Res 2019 Nov 20;21(11):e16197 [FREE Full text] [doi: 10.2196/16197] [Medline: 31746771]
53. 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]
54. Olufadi Y. Social networking time use scale (SONTUS): A new instrument for measuring the time spent on the social networking sites. Telemat Inform 2016 May;33(2):452-471. [doi: 10.1016/j.tele.2015.11.002]
55. Ellison NB, Steinfield C, Lampe C. The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. J Comput Mediat Commun 2007;12(1):1143-1168. [doi: 10.1111/j.1083-6101.2007.00367.x]
56. Rosen L, Whaling K, Carrier L, Cheever N, Rokkum J. The Media and Technology Usage and Attitudes Scale: An empirical investigation. Comput Human Behav 2013 Nov 01;29(6):2501-2511 [FREE Full text] [doi: 10.1016/j.chb.2013.06.006] [Medline: 25722534]
57. Sigerson L, Li AY, Cheung MW, Luk JW, Cheng C. Psychometric properties of the Chinese Internet Gaming Disorder Scale. Addict Behav 2017 Nov;74:20-26 [FREE Full text] [doi: 10.1016/j.addbeh.2017.05.031] [Medline: 28558336]
58. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
59. Maindal HT, Sokolowski I, Vedsted P. Translation, adaptation and validation of the American short form Patient Activation Measure (PAM13) in a Danish version. BMC Public Health 2009 Jun 29;9:209 [FREE Full text] [doi: 10.1186/1471-2458-9-209] [Medline: 19563630]
60. 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]
61. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for Patient Engagement: A Systematic Review. Front Psychol 2015 Jan 08;6:2013. [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]
62. DiFilippo KN, Huang W, Andrade JE, Chapman-Novakofski KM. The use of mobile apps to improve nutrition outcomes: A systematic literature review. J Telemed Telecare 2015 Jul;21(5):243-253. [doi: 10.1177/1357633X15572203] [Medline: 25680388]
63. Gandapur Y, Kianoush S, Kelli HM, Misra S, Urrea B, Blaha MJ, et al. The role of mHealth for improving medication adherence in patients with cardiovascular disease: a systematic review. Eur Heart J Qual Care Clin Outcomes 2016 Oct 01;2(4):237-244 [FREE Full text] [doi: 10.1093/ehjqcco/qcw018] [Medline: 29474713]
64. Xiong S, Berkhouse H, Schooler M, Pu W, Sun A, Gong E, et al. Effectiveness of mHealth Interventions in Improving Medication Adherence Among People with Hypertension: a Systematic Review. Curr Hypertens Rep 2018 Aug 07;20(10):86. [doi: 10.1007/s11906-018-0886-7] [Medline: 30088110]
65. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]
66. van den Berg MH, Schoones JW, Vliet Vlieland TPM. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007 Sep 30;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
67. Wu C, Wu Z, Yang L, Zhu W, Zhang M, Zhu Q, et al. Evaluation of the clinical outcomes of telehealth for managing diabetes: A PRISMA-compliant meta-analysis. Medicine (Baltimore) 2018 Oct;97(43):e12962. [doi: 10.1097/MD.0000000000012962] [Medline: 30412116]

68. Greenwood DA, Blozis SA, Young HM, Nesbitt TS, Quinn CC. Overcoming Clinical Inertia: A Randomized Clinical Trial of a Telehealth Remote Monitoring Intervention Using Paired Glucose Testing in Adults With Type 2 Diabetes. J Med Internet Res 2015 Jul 21;17(7):e178-e482 [FREE Full text] [doi: 10.2196/jmir.4112] [Medline: 26199142]
69. Gilling S, Mortz CG, Vestergaard T. Patient Satisfaction and Expectations Regarding Mobile Teledermoscopy in General Practice for Diagnosis of Non-melanoma Skin Cancer and Malignant Melanoma. Acta Derm Venereol 2020 Apr 21;100(8):adv00117 [FREE Full text] [doi: 10.2340/00015555-3459] [Medline: 32179927]
70. Hawthorne G, Sansoni J, Hayes L, Marosszeky N, Sansoni E. Measuring patient satisfaction with health care treatment using the Short Assessment of Patient Satisfaction measure delivered superior and robust satisfaction estimates. J Clin Epidemiol 2014 May;67(5):527-537. [doi: 10.1016/j.jclinepi.2013.12.010] [Medline: 24698296]
71. Risser NL. Development of an instrument to measure patient satisfaction with nurses and nursing care in primary care settings. Nurs Res 1975;24(1):45-52. [Medline: 1038021]
72. Marshall G, Hays R. The Patient Satisfaction Questionnaire Short Form (PSQ-18). Santa Monica, CA: RAND Corporation; 1994.
73. Demers L, Weiss-Lambrou R, Ska B. Development of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST). Assist Technol 1996 Jun 30;8(1):3-13. [doi: 10.1080/10400435.1996.10132268] [Medline: 10159726]
74. mHealth Economics 2017 - Current Status and Future Trends in Mobile Health. research2guidance. 2017. URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2020-11-03]
75. Nurmi J, Knittle K, Ginchev T, Khattak F, Helf C, Zwickl P, et al. Engaging Users in the Behavior Change Process With Digitalized Motivational Interviewing and Gamification: Development and Feasibility Testing of the Precious App. JMIR Mhealth Uhealth 2020 Jan 30;8(1):e12884 [FREE Full text] [doi: 10.2196/12884] [Medline: 32003750]
76. Rohde A, Duensing A, Dawczynski C, Godemann J, Lorkowski S, Brombach C. An App to Improve Eating Habits of Adolescents and Young Adults (Challenge to Go): Systematic Development of a Theory-Based and Target Group-Adapted Mobile App Intervention. JMIR Mhealth Uhealth 2019 Aug 12;7(8):e11575 [FREE Full text] [doi: 10.2196/11575] [Medline: 30903746]
77. 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]
78. Hankins M, Fraser A, Hodson A, Hooley C, Smith H. Measuring patient satisfaction for the Quality and Outcomes Framework. Br J Gen Pract 2007 Sep;57(542):737-740 [FREE Full text] [Medline: 17761061]
79. Graham B. Defining and Measuring Patient Satisfaction. J Hand Surg Am 2016 Sep;41(9):929-931. [doi: 10.1016/j.jhsa.2016.07.109] [Medline: 27570227]
80. Batbaatar E, Dorjdagva J, Luvsannyam A, Amenta P. Conceptualisation of patient satisfaction: a systematic narrative literature review. Perspect Public Health 2015 Sep;135(5):243-250. [doi: 10.1177/1757913915594196] [Medline: 26187638]
81. Staniszewska S, Brett J, Mockford C, Barber R. The GRIPP checklist: strengthening the quality of patient and public involvement reporting in research. Int J Technol Assess Health Care 2011 Oct;27(4):391-399. [doi: 10.1017/S0266462311000481] [Medline: 22004782]
82. Domecq JP, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, et al. Patient engagement in research: a systematic review. BMC Health Serv Res 2014 Feb 26;14:89 [FREE Full text] [doi: 10.1186/1472-6963-14-89] [Medline: 24568690]
83. Fergusson D, Monfaredi Z, Pussegoda K, Garritty C, Lyddiatt A, Shea B, et al. The prevalence of patient engagement in published trials: a systematic review. Res Involv Engagem 2018;4:17 [FREE Full text] [doi: 10.1186/s40900-018-0099-x] [Medline: 29796308]
84. Lee DJ, Avulova S, Conwill R, Barocas DA. Patient engagement in the design and execution of urologic oncology research. Urol Oncol 2017 Sep;35(9):552-558. [doi: 10.1016/j.urolonc.2017.07.002] [Medline: 28755961]
85. Sheridan S, Schrandt S, Forsythe L, Hilliard TS, Paez KA, Advisory Panel on Patient Engagement (2013 inaugural panel). The PCORI Engagement Rubric: Promising Practices for Partnering in Research. Ann Fam Med 2017 Mar;15(2):165-170 [FREE Full text] [doi: 10.1370/afm.2042] [Medline: 28289118]
86. Forsythe LP, Carman KL, Szydlowski V, Fayish L, Davidson L, Hickam DH, et al. Patient Engagement In Research: Early Findings From The Patient-Centered Outcomes Research Institute. Health Aff (Millwood) 2019 Mar;38(3):359-367. [doi: 10.1377/hlthaff.2018.05067] [Medline: 30830822]
87. Hemphill R, Forsythe LP, Heckert AL, Amolegbe A, Maurer M, Carman KL, et al. What motivates patients and caregivers to engage in health research and how engagement affects their lives: Qualitative survey findings. Health Expect 2020 Apr;23(2):328-336 [FREE Full text] [doi: 10.1111/hex.12979] [Medline: 31800154]
88. Batalden M, Batalden P, Margolis P, Seid M, Armstrong G, Opipari-Arrigan L, et al. Coproduction of healthcare service. BMJ Qual Saf 2016 Jul 16;25(7):509-517 [FREE Full text] [doi: 10.1136/bmjqs-2015-004315] [Medline: 26376674]
89. Radl-Karimi C, Nicolaisen A, Sodemann M, Batalden P, von Plessen C. Coproduction of healthcare service with immigrant patients: protocol of a scoping review. BMJ Open 2018 Feb 03;8(2):e019519 [FREE Full text] [doi: 10.1136/bmjopen-2017-019519] [Medline: 29431137]

90. Policy for device software functions and mobile medical applications. US Food and Drug Administration. 2019 Sep 26. URL: http://www.fda.gov/regulatory-information/search-fda-guidance-documents/policy-device-software-functions-and-mobile-medical-applications [accessed 2020-11-03]
91. Public COTGPOMH. Shaping Europe's digital future - Policy- mHealth. European Commission. 2017 May 09. URL: https://ec.europa.eu/digital-single-market/en/mhealth [accessed 2020-11-03]
92. Developing an approach for using health technology assessments in reimbursement systems for medical products. World Health Organization. 2015 Oct. URL: http://www.who.int/medicines/areas/access/WHO_EMP_PAU_2015_5/en/ [accessed 2020-11-03]
93. Shaping Europe's digital future: Call for Expression of Interesttablishing a working group on mHealth assessment guidelines. European Commission. 2015 Oct 30. URL: https://ec.europa.eu/digital-single-market/en/news/call-expression-interest-establishing-working-group-mhealth-assessment-guidelines [accessed 2020-11-03]
94. International Patient Decision Aids Standards (IPDAS) Collaboration. URL: http://ipdas.ohri.ca/ [accessed 2020-11-03]
95. Ghafur S, Schneider EC. Engaging Patients Using Digital Technology - Learning From Other Industries. NEJM Catalyst. 2019 Jun 05. URL: https://catalyst.nejm.org/patients-digital-consumer-focused-industries/ [accessed 2020-11-03]

Abbreviations

DBCI: digital behavior change intervention
IOM: Institute of Medicine
IPDAS: International Patient Decision Aids
mHealth: mobile health

Edited by G Eysenbach, L Buis; submitted 20.12.19; peer-reviewed by R Buie, V Yim; comments to author 23.03.20; revised version received 11.04.20; accepted 23.10.20; published 11.11.20.

Please cite as:
Bruce C, Harrison P, Giammattei C, Desai SN, Sol JR, Jones S, Schwartz R
Evaluating Patient-Centered Mobile Health Technologies: Definitions, Methodologies, and Outcomes
JMIR Mhealth Uhealth 2020;8(11):e17577
URL: http://mhealth.jmir.org/2020/11/e17577/
doi:10.2196/17577
PMID:33174846

Original Paper

Scientific Publication Patterns of Mobile Technologies and Apps for Posttraumatic Stress Disorder Treatment: Bibliometric Co-Word Analysis

Atik Kulakli[1*], BA, MSc, MA, PhD; Ivanna Shubina[2*], BA, MA, PhD

[1]Department of Management Information Systems, College of Business Administration, American University of the Middle East, Egaila, Kuwait

[2]Psychology, General Education, Liberal Arts Department, American University of the Middle East, Egaila, Kuwait

[*]all authors contributed equally

Corresponding Author:
Atik Kulakli, BA, MSc, MA, PhD
Department of Management Information Systems
College of Business Administration
American University of the Middle East
Block 6, Building 1
Egaila, 54200
Kuwait
Phone: 965 22251400
Fax: 965 26548484
Email: atik.kulakli@aum.edu.kw

Abstract

Background: Mobile apps are viewed as a promising opportunity to provide support for patients who have posttraumatic stress disorder (PTSD). The development of mobile technologies and apps shows similar trends in PTSD treatment. Therefore, this emerging research field has received substantial attention. Consequently, various research settings are planned for current and further studies.

Objective: The aim of this study was to explore the scientific patterns of research domains related to mobile apps and other technologies for PTSD treatment in scholarly publications, and to suggest further studies for this emerging research field.

Methods: We conducted a bibliometric analysis to identify publication patterns, most important keywords, trends for topicality, and text analysis, along with construction of a word cloud for papers published in the last decade (2010 to 2019). Research questions were formulated based on the relevant literature. In particular, we concentrated on highly ranked sources. Based on the proven bibliometric approach, the data were ultimately retrieved from the Web of Science Core Collection (Clarivate Analytics).

Results: A total of 64 studies were found concerning the research domains. The vast majority of the papers were written in the English language (63/64, 98%) with the remaining article (1/64, 2%) written in French. The articles were written by 323 authors/coauthors from 11 different countries, with the United States predominating, followed by England, Canada, Italy, the Netherlands, Australia, France, Germany, Mexico, Sweden, and Vietnam. The most common publication type was peer-reviewed journal articles (48/64, 75%), followed by reviews (8/64, 13%), meeting abstracts (5/64, 8%), news items (2/64, 3%), and a proceeding (1/64, 2%). There was a mean of 6.4 papers published per year over the study period. There was a 100% increase in the number of publications published from 2016 to 2019 with a mean of 13.33 papers published per year during this latter period.

Conclusions: Although the number of papers on mobile technologies for PTSD was quite low in the early period, there has been an overall increase in this research domain in recent years (2016-2019). Overall, these findings indicate that mobile health tools in combination with traditional treatment for mental disorders among veterans increase the efficiency of health interventions, including reducing PTSD symptoms, improving quality of life, conducting intervention evaluation, and monitoring of improvements. Mobile apps and technologies can be used as supportive tools in managing pain, anger, stress, and sleep disturbance. These findings therefore provide a useful overview of the publication trends on research domains that can inform further studies and highlight potential gaps in this field.

(JMIR Mhealth Uhealth 2020;8(11):e19391) doi:10.2196/19391

KEYWORDS

posttraumatic stress disorder (PTSD); mobile technologies; mobile apps; treatment; text analysis; co-word analysis; bibliometric; Web of Science

Introduction

Mobile technologies have emerged as useful tools to support patients who have posttraumatic stress disorder (PTSD) [1]. Recent studies in this field have mainly focused on the development and evaluation of mobile health (mHealth) tools [2], assessing their efficacy within a health care field [3], and developing a standard for creating PTSD-related mHealth apps [4].

New mobile technologies are accessible and cost-effective tools to assist in identifying PTSD symptoms [5,6] and suicidal behaviors [7]. Mobile technologies are supportive in prognosticating the impact that PTSD symptoms have on treatment efficiency [8], in the evaluation of the effectiveness of psychotherapy [9,10] and its precision [11], and in helping individuals manage their mental health [12,13]. The most common benefits of using mobile interventions are cost-effectiveness, reducing waiting lists, and high accessibility [5].

Thus, mHealth tools have potential for supporting traditional treatment for mental disorders among veterans with PTSD [14]. Mood Tracker [15] and other mHealth apps (eg, Parenting 2GO, PTSD Coach, and PTSD Family Coach) were established for the treatment of military staff diagnosed with PTSD and their families [16], which have been demonstrated to be feasible and relevant. In addition, the efficiency of a mobile app designed to screen trauma-related symptoms as a diagnostic tool for trauma survivors was proven [8].

New strategies for a self-managed and web-based wellness training program for PTSD veterans are recognized as being accessible, low-cost, and efficient to decrease PTSD symptoms [5]. A similar study showed no significant difference in effectiveness between a mindfulness training program delivered by a psychotherapist and the same program provided through a self-directed mobile app [6]. Fraynt et al [17] explored the ability of a mobile app to help support the transition to civilian life among PTSD veterans. Similarly, Pavliscsak et al [18] demonstrated the importance of engagement with an mHealth app among service members with PTSD symptoms in transition to help improve communication skills.

In addition, several studies have examined the use of mobile apps in treatment for smokers with PTSD [19,20] and their use in integrated care [12,13], demonstrating their effectiveness and congruency. A pilot study on the mobile app PTSD Coach demonstrated a significant improvement in quality of life among patients with PTSD [21]. Another study showed the efficiency of the Moodivate app in managing limitations associated with evidence-based psychotherapy by decreasing the symptoms of PTSD and depression among adults [22,23]. Other studies on mobile apps for the treatment of anger symptoms [24,25] and in managing pain among individuals with PTSD [26,27] showed feasibility and therapeutic benefits. Overall, mHealth has been evaluated to be effective in reducing emotional dysregulation among veterans with PTSD [28]. Other comparative studies between in vivo exposure and virtual reality–based exposure therapies for patients with PTSD symptoms indicated that the virtual experience was considered to be a more flexible approach [29,30].

An influential group of studies has focused on how mobile technology can be useful in applying cognitive behavioral therapy (CBT) for various mental health conditions. The special needs of patients with PTSD create a demand for modified approaches other than therapeutic sessions in vivo [31] and more creative tools used in treatment [32]. Recent systematic reviews [9,10,30] have explored the efficacy of the mHealth apps embodying CBT principles for the treatment of various mental disorders, including PTSD. For example, Martinez-Miranda et al [7] assessed the effectiveness of applying a mobile-based app in recognizing individuals diagnosed with PTSD and demonstrating suicidal behaviors. Wang et al [33] indicated the high potential and efficiency of mobile apps in the monitoring and management of mental disorders, including PTSD. Stirman et al [34] assessed an app designed for creating and assessing universal worksheets to help evaluate the accuracy of CBT therapy sessions for patients with PTSD. Evans et al [35] investigated the development of a cognitive assessment tool via mobile technology. A similar study by Price et al [36] suggested that PTSD checklists delivered through a mobile app or on paper were equally efficient.

Therefore, new technology as a psychological tool in treating PTSD has been explored by researchers in various modalities, including mobile apps, mHealth, web-based programs, virtual reality, checklists, and assessments [30]. However, the majority of mHealth apps that are currently available lack clinically validated evidence of their efficacy [3]. Accordingly, the primary aim of this study was to explore scientific publication patterns in the research domain of “mobile technologies and apps” concerning PTSD. We further aimed to reveal the contribution of scientific knowledge by highlighting the gaps and provide new directions of potential development areas for further studies.

Methods

Research Questions

Based on the research scope and objectives, the following four research questions formulated:

> *RQ1: What are the characteristics (descriptive) of the publications? How many papers on “mobile technologies and apps” concerning “posttraumatic stress disorder” have been published between 2010 and 2019?*
>
> *RQ2: Who are the most productive authors/coauthors in this field, and what are their countries of origin? What are the citation metrics of these authors?*

RQ3: In what types of sources are the papers published most frequently? Which organizations mainly contribute to this research area?

RQ4: What are the co-words (keywords/text) associated with these publications?

Bibliometric Study

A bibliometric study enables researchers to explore patterns, trends, associations, and scientific developments related to searched domains, along with interrelated fields over publication data. This analysis requires a structured bibliometric database to obtain appropriate data for answering research questions [37-39].

Bibliometrics is also defined as a statistical method to analyze bibliometric publications data over a wide spectrum such as peer-reviewed journal articles, books, conference proceedings, periodicals, reviews, reports, and related reports. There are various analysis methods for a literature review along with bibliometric tools [37,40-43]. This approach allows for further obtaining more in-depth understanding of a given topic and its publication trends.

In this study, we employed a bibliometric approach for obtaining descriptive publication results, author/coauthor productivity metrics, source impact analysis, along with keyword and most common co-word (text) analysis [38,44-47].

Co-word (Text) Analysis and Word Cloud

A word cloud, also known as a "tag cloud," is a visual representation of text data [39] from various keywords or any given text material [42]. According to the Web of Science dataset structure, a word cloud has four main categories to analyze: abstract, title of the paper, author keywords, and keyword plus (see Multimedia Appendix 1). Depending on the frequency of text data regarding the main categories, the significant terms and tags are highlighted, which are usually single words represented by a single font size and color based on their relative importance. Bold and larger-sized words indicate that the word has more importance and has attracted researchers' increasing attention in the subject domain field [48]. Keywords or any other text datasets among these four categories were collected from the articles to conduct the co-word analysis and to construct a word cloud to illustrate the power of words based on their frequencies in the literature [49].

Data Collection and Extraction

A bibliometric study requires a structured database to analyze publication data. The main two bibliometric databases available for this purpose are ISI Web of Science and Scopus. ISI Web of Science provides data on the highest ranked and impactful (prestigious) sources, whereas Scopus also ranks the same sources in addition to other sources with wider coverage, including conferences, symposia, and congress proceedings. In addition, PubMed is a commonly used database in the medical field. Journals focused on mHealth mostly rank in the highest quartile (Q1) in Web of Science as well. Therefore, we chose to focus on these highly ranked (Q1) and high-impact sources in Web of Science to maintain consistency of citations in a single database with SCI-Expanded and SSCI indexing. We only used PubMed for comparison purposes and as a secondary source of indexing. All databases have their citation count categories. The citation results are significant to keep all publications within the same quartile and for consistency of comparisons at the same level (Q1 ranking and indexing). Finally, publication data were retrieved from the Web of Science database using the search strategy shown in Textbox 1 [9,37,38,40,42].

Textbox 1. Search criteria and strategy.a

TITLE: ("*mobile*" OR "mobile*" OR "*Mobile*" OR "Mobile*") AND
TOPIC: ("PTSD" OR "post-traumatic stress disorder" OR "post traumatic stress disorder" OR "posttraumatic stress disorder") (Journal articles title; keywords, abstract)
Database: PubMed
Database: ISI Web of Science Core Collection (Clarivate Analytics)
Indexes: SCI-EXPANDED, SSCI
Timespan: all years of (2010-2019)

[a]64 were papers found in the Web of Science list, including 40 matching papers in PubMed.

The data were retrieved as plain .txt, .xls, .csv, and .bib file formats for further analysis. The Microsoft Excel and R Language (version R x64 3.6.1) R Studio software with the "bibliometrix" package [39] were used for descriptive and bibliometric data analysis, respectively [48].

Results

Publication Profile and Descriptive Publication Results

A total of 64 publications for the research domain were retrieved from the ISI Web of Science database (Multimedia Appendix 2). In comparison to the PubMed results, 40 papers were found in the same dataset of the Web of Science–retrieved publications. The vast majority of the papers were written in the English language (63/64, 98%) with the remaining article written in French (1/64, 2%). Overall, there were 323 authors/coauthors from 11 different countries, with the United States being the most common, followed by England, Canada, Italy, the Netherlands, Australia, France, Germany, Mexico, Sweden, and Vietnam. The descriptive summary of these publications showed that the majority were peer-reviewed journal articles (48/64, 75%), followed by reviews (8/64, 13%), meeting abstracts (5/64, 8%), news items (2/64, 3), and proceedings (1/64, 2%). There was a mean number of 6.4 papers

published on the topic per year, with a 100% increase in the number of publications found for the period of 2016-2019, along with a corresponding increase in the mean number of papers published per year during this period to 13.33.

Web of Science-Core Collection subject category data were used to categorize the related research domains under the top 12 major subjects, which are summarized in Table 1, comprising topics with at least 3 publications. Other minor subject categories with 2 publications included Computer Science Information Systems and Neurosciences, and those with 1 paper each included Computer Science Interdisciplinary Applications, Education Educational Research, Environmental Sciences, Health Policy Services, Information Science Library Science, Medicine Research Experimental, Nursing, Oncology, Pediatrics, Psychology Experimental, Rehabilitation, and Respiratory System.

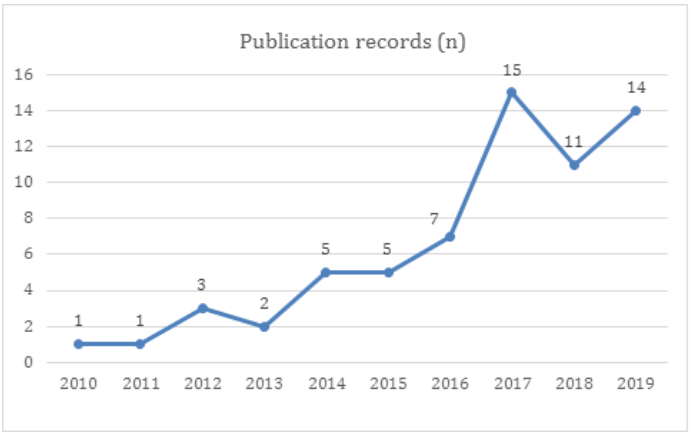
Table 1. Publications in the top Web of Science subject areas from 2010 to 2019 (N=64).

Subject category	Publications, n (%)
Psychiatry	23 (37)
Psychology Clinical	16 (26)
Health Care Sciences Services	11 (18)
Medical Informatics	11 (18)
Medicine General Internal	8 (13)
Psychology Multidisciplinary	6 (10)
Psychology	4 (6)
Substance Abuse	4 (6)
Clinical Neurology	3 (5)
Pharmacology Pharmacy	3 (5)
Public Environmental Occupational Health	3 (5)

Distribution of Publications Over Time

Figure 1 presents the publication records by year starting from 2010 through to the end of the decade in 2019. There was only one publication recorded for 2010 and 2011, followed by a slight increase from 2012 to 2013. From 2013 to 2014, the number of publications increased sharply from 2 to 5, with a steady increase until 2016 (n=5 to n=7). A sharp increase was observed from 2016 (n=7) to 2019 (n=14), representing a 100% increase in the publication count. The last 3 years reflect increasing research interest in the topic, as the period with the highest count in the dataset. The mean number of publications per year for the entire period was 6.4, increasing to 13.33 from 2016 to 2019. The publication distribution also supports this upward movement as shown in Figure 1.

Figure 1. Publication records by year.



Most Productive Authors/Coauthors

Table 2 provides the descriptive results for authors and coauthors of the retrieved publications. Only 6 papers were published as single-authored documents, whereas 317 papers were multiple-authored documents, and there were 323 different authors with 366 appearances.

Table 2. Author and coauthor descriptive results for all documents published between 2010 and 2019.

Descriptive results	Value
Authors	323
Author appearances	366
Authors of single-authored documents	6
Authors of multi-authored documents	317
Documents per author	0.198
Authors per document	2.05
Coauthors per document	5.72
Collaboration index	5.56
Average citations per document	7.95

Among the top 20 most productive authors, Kuhn (n=7 records) was identified as a leading author; followed by Beckham (n=5 records); Calhoun (n=4 records); and Dennis, Marx, and Moore (n=3 records each) in the top category. Others at the same level with 2 papers each include Acierno, Carpenter MJ, Carpenter VL, Dahne, Dedert, Dennis, Diaz, Felton, Hertzberg, Hoffman, Jovanovic, Kirby, and Kizakevich.

Table 3 shows the top 10 most productive countries that contributed to the research domain fields. The top country was the United States, followed by Germany, Italy, the Netherlands, United Kingdom, Australia, Canada, France, Mexico, and Sweden.

Table 3. Most relevant countries by corresponding author.

Country	Articles, N	Frequency	SCP[a]	MCP[b]	MCP/article ratio
USA	45	0.776	42	3	0.07
Germany	2	0.035	2	0	0
Italy	2	0.035	2	0	0
Netherlands	2	0.035	2	0	0
United Kingdom	2	0.035	2	0	0
Australia	1	0.017	1	0	0
Canada	1	0.017	1	0	0
France	1	0.017	1	0	0
Mexico	1	0.017	1	0	0
Sweden	1	0.017	1	0	0

[a]SCP: single-country publication.

[b]MCP: multiple-country publication.

Table 4 shows the contributed institutions from various countries around the world. Harvard University and Stanford University emerged as the leading institutions with 8 papers each, and 13 universities ranked at the top of the dataset among the total 156 contributing institutions.

Table 4. Institutions with the greatest contributions to the field (N=156).

Institution	Publications, n (%)
Harvard University	8 (13)
Stanford University	8 (13)
Duke University	7 (11)
US Department of Veteran Affairs	7 (11)
VA Palo Alto Health Care System	7 (11)
VA Boston Healthcare System	6 (10)
Durham VA Medical Center	5 (8)
Boston University	4 (7)
Harvard Medical School	4 (7)
Medical University of South Carolina	4 (7)
University of California System	4 (7)
University of Washington	4 (7)
University of Washington Seattle	4 (7)

Citation Results

The citation report of the 64 publications derived from the Web of Science Core Collection statistics between 2010 and 2019 showed that the H-index was 12 and the average number of citations per item was 7.9. The sum of times an article was cited in total was 509, which was reduced to 471 after excluding self-citations. The number of cited articles was 421 in total and was 402 after excluding self-citations. Figure 2 shows the total citation distribution throughout the research timeframe. Although the citation statistics are available as of 2012, an increase was only observed as of 2015 (n=28), followed by a steady increase until 2019. From 2016 to 2017 and from 2017 to 2018, the citation counts nearly doubled, reaching 217 by 2019.

Figure 2. Total publication/citation counts per year (2010-2019).

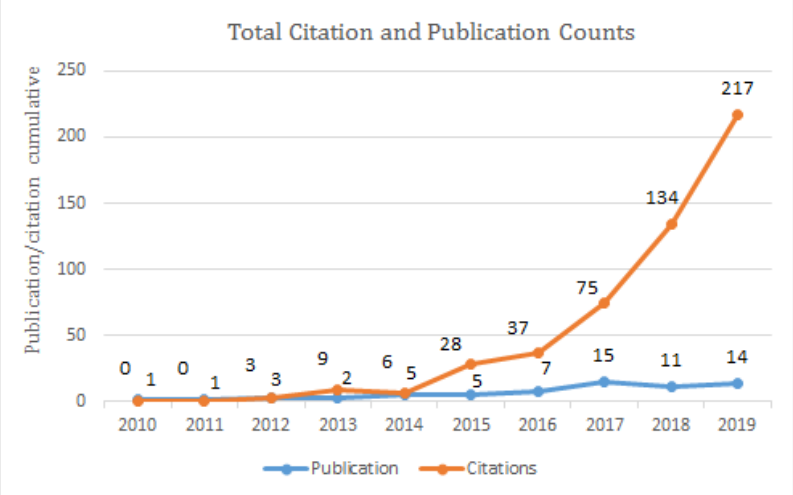


Table 5 presents the top 10 journals by citation counts. Olff [2] was the leading paper with the highest number of citations (n=58) during the search period. The top 10 citations comprised 299 of the total 509 citations (58.7%) starting from 2012 to 2019. There was a mean of 29.9 per year. Therefore, 4 papers were above the mean: Olff [2], Hertzberg et al [19], Van Ameringen et al [50], and Repetto et al [29], published in 2015, 2013, 2017, and 2013, respectively. Although the papers only started to be cited as of 2012, the citation counts showed a substantial (3-times) increase from 2014 to 2015 (from 6 to 20 counts). The most citations were recorded in 2019 (n=108).

The leading journal publishing these studies was *Journal of Medical Internet Research (JMIR)* with 3 articles, and the remaining journals only included 1 article each published in the top 10 list. The majority of these papers were related to mobile technology and apps, PTSD-related topics, virtual reality related to mobile devices, treatment, and other psychological disorders.

Table 5. Top 10 articles by citation results (2010-2019).

Reference	Citations (N)									Average citations/year
	Total	2012	2013	2014	2015	2016	2017	2018	2019	
Olff [2]	58	0	0	0	8	6	17	10	17	9.67
Hertzberg et al [19]	49	0	1	2	8	7	17	8	6	6.13
Van Ameringen et al [50]	43	0	0	0	0	0	2	18	23	10.75
Repetto et al [29]	38	3	8	4	2	6	4	2	9	4.75
Place et al [51]	23	0	0	0	0	0	2	11	10	5.75
Carpenter et al [20]	22	0	0	0	2	4	6	5	5	3.67
Lui et al [3]	21	0	0	0	0	0	0	8	13	5.25
Kahn et al [5]	18	0	0	0	0	0	1	7	10	3.6
Rathbone et al [9]	14	0	0	0	0	0	0	9	5	3.5
Bakker and Rickard [52]	13	0	0	0	0	0	0	3	10	4.33
Sum	299	3	9	6	20	23	49	81	108	

Source Frequency and Productivity

Table 6 shows the publication frequency for the most relevant sources. Although there were no notable differences in publication history, the top contributing journals of 8 sources (n≥2 records) are listed in Table 6.

The top publication sources distributed and sorted from the highest on the list with 4 (19.36%) records were *European Journal of Psychotraumatology*, *JMIR mHealth and uHealth*, and *Military Medicine*. The next most common sources with 3 (14.52%) records were *JMIR*, *Professional Psychology Research and Practice*, and *Psychological Services*, whereas *Applied Psychophysiology and Biofeedback* and *Depression and Anxiety* included 2 records each (6.45%).

There were 39 single-publication sources (n=1, 59.68%) among the total 64 articles in the dataset retrieved: *Addiction*, *Addictive Behaviors*, *American Journal of Preventive Medicine*, *Annales Medico Psychologiques*, *Asian Journal of Psychiatry*, *Behavior Therapy*, *Biological Psychiatry*, *Computers In Human Behavior*, *Current Psychiatry Reports*, *ETR D Educational Technology Research and Development*, *Health Informatics Journal*, *Implementation Science*, *International Journal of Clinical Practice*, *International Journal of Environmental Research and Public Health*, *International Journal of Methods In Psychiatric Research*, *International Journal of Psychiatry In Medicine*, *Jama Journal of The American Medical Association*, *JMIR Medical Informatics*, *JMIR Mental Health*, *Journal of Affective Disorders*, *Journal of Child And Adolescent Psychopharmacology*, *Journal of Clinical Psychiatry*, *Journal of Dual Diagnosis*, *Journal of Head Trauma Rehabilitation*, *Journal of Investigative Medicine*, *Journal of Medical Systems*, *Journal of Pain*, *Journal of Psychiatric Research*, *Journal of Psychosocial Nursing And Mental Health Services*, *Journal of The American Medical Informatics Association*, *Journal of Traumatic Stress*, *Nicotine Tobacco Research*, *Personal and Ubiquitous Computing*, *Psychiatry Interpersonal and Biological Processes*, *Psycho-Oncology*, *Social Psychiatry and Psychiatric Epidemiology*, *Global Mental Health*, *Systematic Reviews*, and *Thorax*.

Table 6. Most relevant top publication sources (N=64).

Source	Records, n (%)
European Journal of Psychotraumatology	4 (6)
JMIR mHealth and uHealth	4 (6)
Military Medicine	4 (6)
Journal of Medical Internet Research (JMIR)	3 (5)
Professional Psychology Research and Practice	3 (5)
Psychological Services	3 (5)
Applied Psychophysiology and Biofeedback	2 (3)
Depression and Anxiety	2 (3)

The *JMIR* group of journals had the highest amount of publication records (n=9) for the research domains in the search period. In addition, the source impact metric (H-index, n≥2) of the journals ranked as follows: *JMIR mHealth and uHealth* (H-index=4); *Military Medicine*, *JMIR*, and *Professional Psychology Research and Practice* (H-index=3); and *European*

Journal of Psychotraumatology, Psychological Services, and *Depression and Anxiety* (H-index=2).

Co-Word (Keyword) Analysis

Keyword analysis is used to reveal the frequency of keywords in publications. There are two types of keywords: author keywords and Keywords Plus. The author keywords are the words (terms) that authors prefer to use for their papers and Keyword Plus includes terms from the preset list of related science domains, which is created by the editorial experts of Web of Science.

Table 7 shows the most relevant keywords used in the publications associated with mobile technologies and apps for PTSD treatment. The most common author keywords included "m-health," "PTSD," "mobile health," and "depression." The most common Keyword Plus terms included "posttraumatic stress disorder," "PTSD," "randomized controlled trial," and "depression."

Table 7. Top 10 keywords.

Rank	Author keywords		Keywords-Plus	
	Term	N	Term	N
1	m-health	14	posttraumatic stress disorder	20
2	PTSD	14	PTSD	13
3	depression	9	randomized controlled trial	13
4	mobile health	7	depression	12
5	mobile apps	6	care	9
6	posttraumatic stress disorder	6	smartphone app	8
7	telemedicine	5	symptoms	8
8	trauma	5	veterans	8
9	mental health	4	meta-analysis	7
10	anxiety	3	prevalence	7

Co-Word (Text) Analysis

The word dynamic-growth graph (Figure 3), prepared with the top keywords, was used to evaluate the keyword dynamics over the research period. The repetition trend of each word (ie, the frequency of appearances in the dataset over the search period) represents occurrences. The graph shows the trend direction to analyze either upward or downward movement over the linear line according to the annual distribution of keywords. The most popular terms and keywords can be tracked during the period to understand trends in subject domain interest and importance in the research field. Identifying topics of growing interest helps researchers to concentrate on new subject areas and can also provide valuable results to contribute to these fields.

Figure 3. Word dynamics-growth.

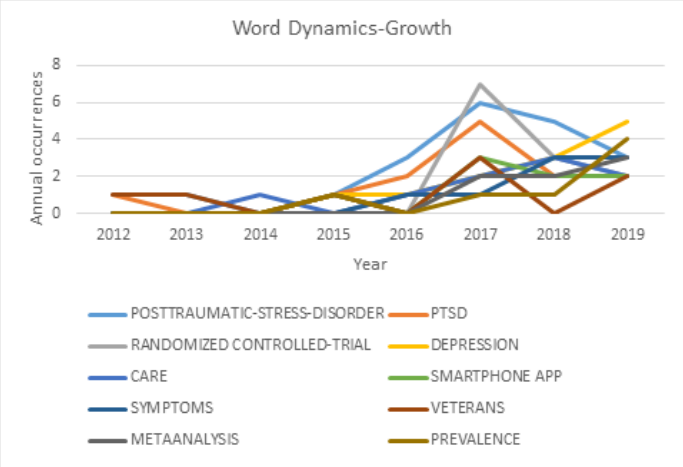


Figure 3 illustrates the word dynamics-growth for the research domains. Although the dataset shows that publications started in 2012, the interest and popularity became significantly more visible after 2015, followed by a sharp increase in the subject areas until the middle of 2017. In those years, "PTSD," "posttraumatic stress disorder," "smartphone," and "randomized controlled trial" had peak levels, but the usage rates of these terms declined from the middle of 2017 to 2019. At this point, the terms "depression" and "prevalence" became highly popular with a sharp increase on the graph, followed by "symptoms" and "meta-analysis."

Multimedia Appendix 1 illustrates four different world cloud–based data distributions. The top left word cloud is for

abstract text data, whereas the upper right word cloud is for author keywords text data. Similarly, the bottom left word cloud is from Keyword Plus text data and the bottom right word cloud shows a paper title text data representation. As can be seen in each word cloud, the highlighted words differ in terms of importance for each group. In particular, "health," "mobile," and "apps" were of the highest importance, followed by "mental," "treatment," and "PTSD" in the abstract word cloud. By contrast, in the author keyword cloud, "m-health" and "PTSD" were both dominant keywords, followed by "depression," "mobile health," and "posttraumatic stress disorder." The word order changed for Keyword Plus to "depression" and "PTSD" at the top, followed by "care," "smartphone app," and "symptoms." The title word cloud was comparable to the others with "mobile" being far more important than other key terms as a highly dominant term, followed by "health," and the terms "stress," "app," "disorder," "mental", "posttraumatic," and "application" appeared in another layer.

Discussion

Principal Results

Mobile apps are recognized as efficient tools in the assistance of both health care patients and staff [1,3]. A mobile platform allows for predicting the symptoms of depression and PTSD [51] to collect data and conduct research on health interventions using the technology [53]. Mobile technology in the treatment for military staff and veterans with PTSD and other mental disorders has been reported to be feasible [8,15]. Engagement with mHealth apps in the transition period improved communications [18] and mental well-being [52]. Several mobile apps have been considered to be useful supportive tools in the treatment of managing pain [26,27], stress reduction [54], anger [25], and mental disorders, including PTSD, depression, anxiety, and addictions [50,55], as well as in identifying suicidal behaviors [7].

The efficiency of mobile apps in the psychological treatment of mental disorders, including CBT [9], cognitive rehabilitation [28], exposure therapy [10,29,30], and visualization [56] has been verified. The use of mobile apps allows for evaluating CBT precision [34], to conduct a cognitive assessment [35], to deliver a PTSD checklist [36], and to monitor mental disorders [33].

By analyzing the results of the dataset regarding the publication pattern on research related to mobile technology and apps associated with PTSD treatment, we found a rapid increase and growth of subject interest in the last decade. The trend sharply increased in the most recent years, from 2016 to 2019.

This growth of productivity could reflect the improvements, functionality, and developments of mobile technology and apps in parallel with comparison to other areas of usage. Therefore, these technologies have become the center of human life to provide new opportunities, convenience, and address potential benefits. Despite the negative insights and perceptions of such technologies, this would be promising for patients to adapt their health behaviors as a supportive tool with their clinical treatments. Cooperation among clinical experts, app designers, and technology providers is necessary to reach the ultimate goal and objectives, which should concentrate on patients' needs and treat them positively and objectively.

This analysis was conducted to highlight the most frequent subject categories, along with popular keywords and terms. These aspects were reflected in keywords, and the same terms were similarly represented in the text of the abstract and title of the paper. The text analysis showed the critical terms used in this field of research and also represents the popularity of subdomain searches. The aim of this study was to discover the publication trend and to identify the critical areas in the dataset to ultimately provide insights and research directions for academics, practitioners, and readers who wish to collaborate in these domains in the future.

According to the papers retrieved and analyzed (N=64), the majority were peer-reviewed journal articles (75%), with a mean of 6.4 publications per year from 2010 to 2016, which then sharply increased to 13.33 (doubled) between 2016 and 2019. The most productive countries were the United States, with far greater representation than any other country, followed by Germany, Italy, the Netherlands, and the United Kingdom. One of the main reasons for this difference is attributed to differences in military engagement in various regions. For example, in the Middle East and Asia, soldiers return to their countries with different intensities of PTSD symptoms and other comorbid disorders.

The results also revealed the distribution of the publications, demonstrating that the top category sources support previous arguments; namely *European Journal of Psychotraumatology*, *JMIR mHealthand uHealth*, and *Military Medicine* were among the highly popular sources for these publications. The *JMIR* group of journals emerged as the leading sources (n=9, 14% coverage) compared to other single-publication sources. However, the vast majority (60%) of sources were equally distributed among the 39 single-publication sources. Similar results were found in the authorship analysis, in which the top contributors in the field are Kuhn (n=7 records) as the leading author; followed by Beckham (n=5 records); Calhoun (n=4 records); and Dennis, Marx, and Moore (n=3 records each) in the top categories.

Co-word (text) analysis showed that the most common vital terms overall (for the four different word clouds created) were "mobile," "PTSD," "posttraumatic stress disorder," "m-health," "depression," "health," "treatment," "smartphone apps," and "mobile health."

Strengths and Limitations

Despite growing interest for the research domains, no publication was identified that analyzed the state of the field with a bibliometric approach. Therefore, the main strength of this study could be considered as the uniqueness of the research design itself. This study is the first bibliometric-related research in the domain. The contribution of the study is revealing the scientific patterns and future research gaps to academics and practitioners. The text analysis also highlighted and supported popular subject areas to clarify the research scope and future directions.

One of the limitations of this study is that we used only the Web of Science Core Collection database in comparison to PubMed. A single database was selected to ensure a simple and accurate analysis, and to effectively eliminate duplications and avoid errors. In this regard, Web of Science covers the highest impact journals and has unique indexing and ranking with its citation categories. The research domain of health and mobile internet–related publications is ranked in the SCI-Exp index. Another limitation is that only documents published in the English language were selected. Although various bibliometric analysis methods are available, given the scope and size of this topic, we decided to concentrate on more specific analyses such as descriptive statistics regarding the dataset from 2010 to 2019.

Future Research Suggestions

According to the findings, the research domains are prevalent, and growing interest can be seen as an upward trend in the publication records since 2016. In particular, the majority of the subject category records were found in Psychiatry and Psychology, especially in the clinical and multidisciplinary domains, followed by Health Care Science Services, Medical Informatics, and Medicine General Internal. Further research is needed adopting various aspects of bibliometric analysis. More empirical and case studies should also be conducted in parallel with the improvement of technology and apps perspectives that would be tested and clinically validated.

However, the research analyzed indicates the importance of further explorations to develop appropriate and feasible mobile technology for PTSD treatment. The necessity to manage the challenges related to the development of mHealth tools were underlined [2]. Establishing the standards for creating PTSD-related mHealth apps and following them seem to be essential in transferring mobile apps to the clinical field [4]. Moreover, practitioners need to explore the factors facilitating and limiting the effective use of mHealth for PTSD treatment [57,58]. Cooperation between mobile app creators, researchers, and practitioners is essential in creating new technology that will match the needs and expectations of both health care staff and patients [50]. Finally, the majority of the available mobile apps require more clinically validated evidence of their efficacy before they can be adopted in the psychological treatment of PTSD [33].

Conclusions

This study explored and analyzed the scientific patterns and relations of scholarly publications related to the use of mobile technologies in PTSD treatment. We therefore provide a general overview of the field based on co-word (text and keyword) analysis of research domains, and various forms of bibliometric methods were employed along with a data visualization approach to establish a clear picture. The analysis included 64 papers published between 2010 and 2019.

The results identify the most frequent subject categories, popular keywords, critical terms, and the popularity of subdomain searches. With this study, we attempted to investigate the patterns of publications to provide insights and research directions for academics, practitioners, and readers who wish to collaborate in these domains in the future. The data highlight the significance of further explorations in this field to improve mobile technology for PTSD treatment. Conducting studies and analyzing the practical use of these tools will improve the technology and apps that would be tested and clinically validated.

Authors' Contributions

AK designed the methodology and performed data extraction. IS performed the literature review (background). Both authors conceived of the study, and drafted, reviewed, and edited the manuscript. Both authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Word clouds visualization.
[PNG File , 230 KB - mhealth_v8i11e19391_app1.png]

Multimedia Appendix 2
Repository dataset of publications.
[PDF File (Adobe PDF File), 490 KB - mhealth_v8i11e19391_app2.pdf]

References

1. Rodriguez-Paras C, Tippey K, Brown E, Sasangohar F, Creech S, Kum H, et al. Posttraumatic Stress Disorder and Mobile Health: App Investigation and Scoping Literature Review. JMIR Mhealth Uhealth 2017 Oct 26;5(10):e156 [FREE Full text] [doi: 10.2196/mhealth.7318] [Medline: 29074470]
2. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015;6:27882 [FREE Full text] [doi: 10.3402/ejpt.v6.27882] [Medline: 25994025]
3. Lui JHL, Marcus DK, Barry CT. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. Prof Psychol Res Pract 2017 Jun;48(3):199-210. [doi: 10.1037/pro0000122]

4. Schellong J, Lorenz P, Weidner K. Proposing a standardized, step-by-step model for creating post-traumatic stress disorder (PTSD) related mobile mental health apps in a framework based on technical and medical norms. Eur J Psychotraumatol 2019;10(1):1611090 [FREE Full text] [doi: 10.1080/20008198.2019.1611090] [Medline: 31143412]
5. Kahn JR, Collinge W, Soltysik R. Post-9/11 Veterans and Their Partners Improve Mental Health Outcomes with a Self-directed Mobile and Web-based Wellness Training Program: A Randomized Controlled Trial. J Med Internet Res 2016 Sep 27;18(9):e255 [FREE Full text] [doi: 10.2196/jmir.5800] [Medline: 27678169]
6. Cox CE, Hough CL, Jones DM, Ungar A, Reagan W, Key MD, et al. Effects of mindfulness training programmes delivered by a self-directed mobile app and by telephone compared with an education programme for survivors of critical illness: a pilot randomised clinical trial. Thorax 2019 Jan;74(1):33-42 [FREE Full text] [doi: 10.1136/thoraxjnl-2017-211264] [Medline: 29793970]
7. Martínez-Miranda J, Martínez A, Ramos R, Aguilar H, Jiménez L, Arias H, et al. Assessment of users' acceptability of a mobile-based embodied conversational agent for the prevention and detection of suicidal behaviour. J Med Syst 2019 Jun 25;43(8):246. [doi: 10.1007/s10916-019-1387-1] [Medline: 31240494]
8. van der Meer CAI, Bakker A, Schrieken BAL, Hoofwijk MC, Olff M. Screening for trauma-related symptoms via a smartphone app: The validity of Smart Assessment on your Mobile in referred police officers. Int J Methods Psychiatr Res 2017 Sep;26(3):e1579 [FREE Full text] [doi: 10.1002/mpr.1579] [Medline: 28948699]
9. Rathbone AL, Clarry L, Prescott J. Assessing the Efficacy of Mobile Health Apps Using the Basic Principles of Cognitive Behavioral Therapy: Systematic Review. J Med Internet Res 2017 Nov 28;19(11):e399 [FREE Full text] [doi: 10.2196/jmir.8598] [Medline: 29187342]
10. Whiteside SPH. Mobile Device-Based Applications for Childhood Anxiety Disorders. J Child Adolesc Psychopharmacol 2016 Apr;26(3):246-251. [doi: 10.1089/cap.2015.0010] [Medline: 26244903]
11. Resick PA, Bovin MJ, Calloway AL, Dick AM, King MW, Mitchell KS, et al. A critical evaluation of the complex PTSD literature: implications for DSM-5. J Trauma Stress 2012 Jun;25(3):241-251. [doi: 10.1002/jts.21699] [Medline: 22729974]
12. Hicks TA, Thomas SP, Wilson SM, Calhoun PS, Kuhn ER, Beckham JC. A Preliminary Investigation of a Relapse Prevention Mobile Application to Maintain Smoking Abstinence Among Individuals With Posttraumatic Stress Disorder. J Dual Diagn 2017;13(1):15-20 [FREE Full text] [doi: 10.1080/15504263.2016.1267828] [Medline: 27918881]
13. Herbst E, Pennington D, Kuhn E, McCaslin SE, Delucchi K, Batki SL, et al. Mobile Technology for Treatment Augmentation in Veteran Smokers With Posttraumatic Stress Disorder. Am J Prev Med 2018 Jan;54(1):124-128. [doi: 10.1016/j.amepre.2017.08.016] [Medline: 29074319]
14. Smith B, Harms WD, Burres S, Korda H, Rosen H, Davis J. Enhancing behavioral health treatment and crisis management through mobile ecological momentary assessment and SMS messaging. Health Informatics J 2012 Dec;18(4):294-308 [FREE Full text] [doi: 10.1177/1460458212445349] [Medline: 23257059]
15. Bush NE, Ouellette G, Kinn J. Utility of the T2 Mood Tracker mobile application among army warrior transition unit service members. Mil Med 2014 Dec;179(12):1453-1457. [doi: 10.7205/MILMED-D-14-00271] [Medline: 25469967]
16. Nolan J, Lindeman S, Varghese FP. Mobile app interventions for military and veteran families: Before, during, and after deployment. Psychol Serv 2019 May;16(2):208-212. [doi: 10.1037/ser0000272] [Medline: 30407049]
17. Fraynt R, Cooper D, Edwards-Stewart A, Hoyt T, Micheel L, Pruitt L, et al. An evaluation of mobile applications designed to assist service members and veterans transitioning to civilian life. Psychol Serv 2018 May;15(2):208-215. [doi: 10.1037/ser0000205] [Medline: 29723023]
18. Pavliscsak H, Little JR, Poropatich RK, McVeigh FL, Tong J, Tillman JS, et al. Assessment of patient engagement with a mobile application among service members in transition. J Am Med Inform Assoc 2016 Jan;23(1):110-118. [doi: 10.1093/jamia/ocv121] [Medline: 26363048]
19. Hertzberg JS, Carpenter VL, Kirby AC, Calhoun PS, Moore SD, Dennis MF, et al. Mobile contingency management as an adjunctive smoking cessation treatment for smokers with posttraumatic stress disorder. Nicotine Tob Res 2013 Nov;15(11):1934-1938. [doi: 10.1093/ntr/ntt060] [Medline: 23645606]
20. Carpenter VL, Hertzberg JS, Kirby AC, Calhoun PS, Moore SD, Dennis MF, et al. Multicomponent smoking cessation treatment including mobile contingency management in homeless veterans. J Clin Psychiatry 2015 Jul;76(7):959-964 [FREE Full text] [doi: 10.4088/JCP.14m09053] [Medline: 25699616]
21. Tiet QQ, Duong H, Davis L, French R, Smith CL, Leyva YE, et al. PTSD coach mobile application with brief telephone support: A pilot study. Psychol Serv 2019 May;16(2):227-232. [doi: 10.1037/ser0000245] [Medline: 30407056]
22. Dahne J, Lejuez CW, Kustanowitz J, Felton JW, Diaz VA, Player MS, et al. Moodivate: A self-help behavioral activation mobile app for utilization in primary care-Development and clinical considerations. Int J Psychiatry Med 2017 Mar;52(2):160-175 [FREE Full text] [doi: 10.1177/0091217417720899] [Medline: 28792292]
23. Dahne J, Lejuez CW, Diaz VA, Player MS, Kustanowitz J, Felton JW, et al. Pilot Randomized Trial of a Self-Help Behavioral Activation Mobile App for Utilization in Primary Care. Behav Ther 2019 Jul;50(4):817-827 [FREE Full text] [doi: 10.1016/j.beth.2018.12.003] [Medline: 31208690]
24. Mackintosh M, Niehaus J, Taft CT, Marx BP, Grubbs K, Morland LA. Using a Mobile Application in the Treatment of Dysregulated Anger Among Veterans. Mil Med 2017 Nov;182(11):e1941-e1949. [doi: 10.7205/MILMED-D-17-00063] [Medline: 29087863]

25. Morland LA, Niehaus J, Taft C, Marx BP, Menez U, Mackintosh M. Using a Mobile Application in the Management of Anger Problems Among Veterans: A Pilot Study. Mil Med 2016 Sep;181(9):990-995. [doi: 10.7205/MILMED-D-15-00293] [Medline: 27612342]
26. Johnson SS, Levesque DA, Broderick LE, Bailey DG, Kerns RD. Pain Self-Management for Veterans: Development and Pilot Test of a Stage-Based Mobile-Optimized Intervention. JMIR Med Inform 2017 Oct 17;5(4):e40 [FREE Full text] [doi: 10.2196/medinform.7117] [Medline: 29042341]
27. Pacella M, Germain A, Suffoletto B, Kuhn E, Jaramillo S, Callaway C. (148) A Randomized Controlled Trial of the PTSD Coach Mobile Health App at Reducing Pain and Psychological Symptoms among Injured Emergency Department Patients: Preliminary Results. J Pain 2019 Apr;20(4):S13. [doi: 10.1016/j.jpain.2019.01.067]
28. Elbogen EB, Dennis PA, Van Voorhees EE, Blakey SM, Johnson JL, Johnson SC, et al. Cognitive Rehabilitation With Mobile Technology and Social Support for Veterans With TBI and PTSD: A Randomized Clinical Trial. J Head Trauma Rehabil 2019;34(1):1-10 [FREE Full text] [doi: 10.1097/HTR.0000000000000435] [Medline: 30169439]
29. Repetto C, Gaggioli A, Pallavicini F, Cipresso P, Raspelli S, Riva G. Virtual reality and mobile phones in the treatment of generalized anxiety disorders: a phase-2 clinical trial. Pers Ubiquit Comput 2011 Oct 2;17(2):253-260. [doi: 10.1007/s00779-011-0467-0]
30. Shubina I. Analysis of the outcomes for the treatment of post-traumatic stress disorder among children and adolescent: systematic review. Int J Psychosoc Rehab. 2020 May. URL: https://www.psychosocial.com/article/PR290320/26205/ [accessed 2020-08-15] [WebCite Cache ID https://www.psychosocial.com/article/PR290320/26205/]
31. Shubina I. Cognitive-behavioral Therapy of Patients with Ptsd: Literature Review. Procedia Soc Behav Sci 2015 Jan;165:208-216. [doi: 10.1016/j.sbspro.2014.12.624]
32. Shubina I. Creativity in psychotherapy: the possibilities of its utilization. In: The European Proceeding of Social & Behavioural Sciences. 2017 Apr 26 Presented at: The International Congress on Clinical & Counselling Psychology (CPSYC) 2017; April 26-28, 2017; Bari, Italy p. 99-111. [doi: 10.15405/epsbs.2017.05.12]
33. Wang K, Varma DS, Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. J Psychiatr Res 2018 Dec;107:73-78. [doi: 10.1016/j.jpsychires.2018.10.006] [Medline: 30347316]
34. Wiltsey Stirman S, Marques L, Creed TA, Gutner CA, DeRubeis R, Barnett PG, et al. Leveraging routine clinical materials and mobile technology to assess CBT fidelity: the Innovative Methods to Assess Psychotherapy Practices (imAPP) study. Implement Sci 2018 May 22;13(1):69 [FREE Full text] [doi: 10.1186/s13012-018-0756-3] [Medline: 29789017]
35. Evans J, Russoniello C, Pougatchev V, Bart-Knauer B, Murray N, Fish M, et al. Development of a Mobile Naturalistic Cognitive Assessment. Applied Psychophysiology and Biofeedback 39(3-4) 2014 Sep 18;39(3-4):313. [doi: 10.1007/s10484-014-9254-9]
36. Price M, Kuhn E, Hoffman JE, Ruzek J, Acierno R. Comparison of the PTSD Checklist (PCL) Administered via a Mobile Device Relative to a Paper Form. J Trauma Stress 2015 Oct;28(5):480-483. [doi: 10.1002/jts.22037] [Medline: 26375277]
37. Kulakli A, Osmanaj V. Global Research on Big Data in Relation with Artificial Intelligence (A Bibliometric Study: 2008-2019). Int J Onl Eng 2020 Feb 12;16(02):31-46. [doi: 10.3991/ijoe.v16i02.12617]
38. Kulakli A, Shubina I. A Bibliometric Study on Mobile Applications for PTSD Treatment: The Period of 2010-2019. : IEEE; 2020 Mar 27 Presented at: 6th International Conference on Information Management (ICIM); March 27-29, 2020; London, UK p. 314-318 URL: https://ieeexplore.ieee.org/document/9081389 [doi: 10.1109/icim49319.2020.244717]
39. Aria M, Cuccurullo C. bibliometrix : An R-tool for comprehensive science mapping analysis. J Informetr 2017 Nov;11(4):959-975. [doi: 10.1016/j.joi.2017.08.007]
40. Jalali MS, Razak S, Gordon W, Perakslis E, Madnick S. Health Care and Cybersecurity: Bibliometric Analysis of the Literature. J Med Internet Res 2019 Feb 15;21(2):e12644 [FREE Full text] [doi: 10.2196/12644] [Medline: 30767908]
41. 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 [FREE Full text] [doi: 10.2196/jmir.8954] [Medline: 29669703]
42. Li F, Li M, Guan P, Ma S, Cui L. Mapping publication trends and identifying hot spots of research on Internet health information seeking behavior: a quantitative and co-word biclustering analysis. J Med Internet Res 2015 Mar 25;17(3):e81 [FREE Full text] [doi: 10.2196/jmir.3326] [Medline: 25830358]
43. Cui L. Rating health web sites using the principles of citation analysis: a bibliometric approach. J Med Internet Res 1999;1(1):E4 [FREE Full text] [doi: 10.2196/jmir.1.1.e4] [Medline: 11720913]
44. Taiebi Javid E, Nazari M, Ghaeli MR. Social media and e-commerce: A scientometrics analysis. Int J Data Netw Sci 2019;3(3):269-290. [doi: 10.5267/j.ijdns.2019.2.001]
45. Gilani E, Salimia D, Jouyandeh M, Tavasoli K, Wong W. A trend study on the impact of social media in decision making. Int J Data Netw Sci 2019;3(3):201-222. [doi: 10.5267/j.ijdns.2019.2.004]
46. Jelvehgaran Esfahani H, Tavasoli K, Jabbarzadeh A. Big data and social media: A scientometrics analysis. Int J Data Netw Sci 2019;3(3):145-164. [doi: 10.5267/j.ijdns.2019.2.007]

47. Tran BX, Latkin CA, Sharafeldin N, Nguyen K, Vu GT, Tam WWS, et al. Characterizing Artificial Intelligence Applications in Cancer Research: A Latent Dirichlet Allocation Analysis. JMIR Med Inform 2019 Sep 15;7(4):e14401 [FREE Full text] [doi: 10.2196/14401] [Medline: 31573929]
48. Birgun S, Kulakli A. Scientific publication analysis on lean management in healthcare sector: the period of 2010-2019. Istanbul Ticaret Universitesi Sosyal Bilimler Dergisi, 19 (Temmuz 2020 Ozel Ek). Istanbul Turkey: Istanbul Ticaret Universitesi Sosyal Bilimler Dergisi; 2020 Jul 31. URL: https://dergipark.org.tr/en/pub/iticusbe/issue/56181/769294 [accessed 2020-08-15] [WebCite Cache ID https://dergipark.org.tr/en/pub/iticusbe/issue/56181/769294]
49. Shen L, Xiong B, Hu J. Research status, hotspots and trends for information behavior in China using bibliometric and co-word analysis. J Document 2017 Jul 10;73(4):618-633. [doi: 10.1108/jd-10-2016-0125]
50. 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]
51. Place S, Blanch-Hartigan D, Rubin C, Gorrostieta C, Mead C, Kane J, et al. Behavioral Indicators on a Mobile Sensing Platform Predict Clinically Validated Psychiatric Symptoms of Mood and Anxiety Disorders. J Med Internet Res 2017 Mar 16;19(3):e75 [FREE Full text] [doi: 10.2196/jmir.6678] [Medline: 28302595]
52. Bakker D, Rickard N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. J Affect Disord 2018 Feb;227:432-442. [doi: 10.1016/j.jad.2017.11.016] [Medline: 29154165]
53. Eckhoff RP, Kizakevich PN, Bakalov V, Zhang Y, Bryant SP, Hobbs MA. A Platform to Build Mobile Health Apps: The Personal Health Intervention Toolkit (PHIT). JMIR Mhealth Uhealth 2015 Jun 01;3(2):e46 [FREE Full text] [doi: 10.2196/mhealth.4202] [Medline: 26033047]
54. Kizakevich PN, Eckhoff R, Brown J, Tueller SJ, Weimer B, Bell S, et al. PHIT for Duty, a Mobile Application for Stress Reduction, Sleep Improvement, and Alcohol Moderation. Mil Med 2018 Mar 01;183(suppl_1):353-363. [doi: 10.1093/milmed/usx157] [Medline: 29635566]
55. Turvey C, Fortney J. The Use of Telemedicine and Mobile Technology to Promote Population Health and Population Management for Psychiatric Disorders. Curr Psychiatry Rep 2017 Oct 16;19(11):88. [doi: 10.1007/s11920-017-0844-0] [Medline: 29035422]
56. Chittaro L, Sioni R. Evaluating mobile apps for breathing training: The effectiveness of visualization. Comput Hum Behav 2014 Nov;40:56-63. [doi: 10.1016/j.chb.2014.07.049]
57. Reger GM, Browne KC, Campellone TR, Simons C, Kuhn E, Fortney JC, et al. Barriers and facilitators to mobile application use during PTSD treatment: Clinician adoption of PE coach. Prof Psychol Res Pract 2017 Dec;48(6):510-517. [doi: 10.1037/pro0000153] [Medline: 28425725]
58. Lipschitz J, Miller CJ, Hogan TP, Burdick KE, Lippin-Foster R, Simon SR, et al. Adoption of Mobile Apps for Depression and Anxiety: Cross-Sectional Survey Study on Patient Interest and Barriers to Engagement. JMIR Ment Health 2019 Jan 25;6(1):e11334 [FREE Full text] [doi: 10.2196/11334] [Medline: 30681968]

Abbreviations

CBT: cognitive behavioral therapy
mHealth: mobile health
PTSD: posttraumatic stress disorder

Edited by T Rashid Soron; submitted 15.04.20; peer-reviewed by A Ferrario, ZE Garcia-Batista, A Sudaryanto, S Birgün; comments to author 15.05.20; revised version received 31.08.20; accepted 13.10.20; published 26.11.20.

Please cite as:
Kulakli A, Shubina I
Scientific Publication Patterns of Mobile Technologies and Apps for Posttraumatic Stress Disorder Treatment: Bibliometric Co-Word Analysis
JMIR Mhealth Uhealth 2020;8(11):e19391
URL: http://mhealth.jmir.org/2020/11/e19391/
doi:10.2196/19391
PMID:33242019

Review

Nurses' Use of Personal Smartphone Technology in the Workplace: Scoping Review

Andrea de Jong[1*], BSCN, MSCN; Lorie Donelle[1*], BSCN, PhD; Michael Kerr[1*], PhD

Arthur Labatt Family School of Nursing, Faculty of Health Sciences, Western University, London, ON, Canada

[*]all authors contributed equally

Corresponding Author:
Lorie Donelle, BSCN, PhD
Arthur Labatt Family School of Nursing
Faculty of Health Sciences
Western University
FNB Rm 2356 Faculty of Health Science Western Uni
1115 Richmond St
London, ON, N6A 5B9
Canada
Phone: 1 5196612111 ext 86565
Email: ldonelle@uwo.ca

Abstract

Background: There has been an increase in the technological infrastructures of many health care organizations to support the practice of health care providers. However, many nurses are using their personal digital devices, such as smartphones, while at work for personal and professional purposes. Despite the proliferation of smartphone use in the health care setting, there is limited research on the clinical use of these devices by nurses. It is unclear as to what extent and for what reasons nurses are using their personal smartphones to support their practice.

Objective: This review aimed to understand the current breadth of research on nurses' personal smartphone use in the workplace and to identify implications for research, practice, and education.

Methods: A scoping review using Arksey and O'Malley's methodological framework was conducted, and the following databases were used in the literature search: CINAHL, PubMed, ProQuest Dissertations and Theses, Embase, MEDLINE, Nursing and Allied Health Database, Scopus, Web of Science, and Cochrane Reviews. Search terms used were Nurs* AND (personal digital technology OR smartphone OR cellphone OR mobile phone OR cellular phone). Inclusion criteria included research focused on nurses' use of their own digital technologies, reported in English, and published between January 2010 and January 2020. Exclusion criteria were if the device or app was implemented for research purposes, if it was provided by the organization, if it focused on infection control, and if it was focused on nursing students or nursing education.

Results: A total of 22 out of 2606 articles met the inclusion criteria. Two main themes from the thematic analyses included *personal smartphone use for patient care* and *implications of personal smartphone use*. Nurses used their smartphones to locate information about medications, procedures, diagnoses, and laboratory tests. Downloaded apps were used by nurses to locate patient care–related information. Nurses reported improved communication among health team members and used their personal devices to communicate patient information via text messaging, calling, and picture and video functions. Nurses expressed insight into personal smartphone use and challenges related to distraction, information privacy, organizational policies, and patient perception.

Conclusions: Nurses view personal smartphones as an efficient method to gather patient care information and to communicate with the health care team. This review highlights knowledge gaps regarding nurses' personal device use and information safety, patient care outcomes, and communication practices. This scoping review facilitates critical reflection on patient care practices within the digital context. We infer that nurses' use of their personal devices to communicate among the health care team may demonstrate a technological "work-around" meant to reconcile health system demands for cost-efficiency with efforts to provide quality patient care. The current breadth of research is focused on acute care, with little research focus in other practices settings. Research initiatives are needed to explore personal device use across the continuum of health care settings.

(JMIR Mhealth Uhealth 2020;8(11):e18774) doi:10.2196/18774

KEYWORDS

nurses; digital health; smartphone; evidence-informed practice

Introduction

Smartphone use is increasing globally, with an estimated 3.3 billion users in 2019 and an anticipated 3.8 billion users by 2021 [1]. Countries with the greatest number of smartphone users include China, India, and the United States [1]. A smartphone is a device that has both computing abilities and mobile communication technology [2]. Mobile communication technology has undergone rapid development [3], and smartphone production is one of the fastest growing technological industries [4]. Coupled with the proliferation of smartphones is the high demand for downloadable apps that a smartphone user can add to their device. In 2018, over 4 million Android and Apple apps were available either free or for a cost [5]. These apps provide the user with additional services and features on their smartphone device.

Within the health care context, there has been an increase in the use of smartphones by health care providers in the last several years [2], including an increase in the use of health professional apps [6]. Fölster [7] reported over 200,000 health-related apps available for purchase or free of charge. The increased use of smartphones and selected apps by health care professionals mirrors the increased use of information technology within health care. There has been a significant increase in the implementation and use of electronic documentation systems, with billions of dollars being utilized to support the provision and coordination of health care to patients using information technology [8]. However, despite the increased use of apps in health care, health care professionals have expressed concerns relating to the trustworthiness of apps and a knowledge gap on the effectiveness of apps [9].

Electronic documentation systems have been integrated within health care systems to provide a centralized repository and accessible source of clinical information for health care providers [10]. These same systems are intended to support clinical care practices among all health care providers while providing access to information, decision-support tools, and improved workflow [10,11]. Given their extended contact with patients and families, nurses tend to have the greatest interaction with electronic documentation systems relative to other health care providers [12].

Despite the increase in technological infrastructures of many health care organizations to support the practice of health care providers, many nurses are using their personal digital devices (eg, smartphones) while at work for both professional and personal purposes [13,14]. Yet, there is limited research available on the clinical use of personal smartphone devices by nurses. In addition, there is limited research available that focuses on nurses' use of commercially available health-related apps in the workplace [15]. These gaps in research are especially important to understand given the most recent global pandemic caused by COVID-19 [16]. The COVID-19 pandemic lead to rapidly evolving public health measures with frequent changes in practice guidelines in all health care settings. Health care organizations had to respond and adapt quickly. With smartphones being a convenient and accessible way to locate information, the use of these devices in practice by nurses needs to be an area of focus, especially when facing a global pandemic.

It is unclear as to what extent and for what reasons nurses are using their personal smartphones to support their practice. Similarly, there is a lack of clarity regarding the quality and credibility of the resources accessed by nurses via their personal devices. However, the expectation of evidence-informed patient care practices highlights the importance of access to health-related resources for nurses and other health care providers [17]. In essence, evidence-informed practice necessitates nurses to critically evaluate information collected regarding patients' needs and to integrate it with the available clinical and research evidence to enable evidence-informed patient care [18]. A valued feature of mobile smartphones is the *pro re nata*, or PRN, access to vast amounts of online health-related information. In this scoping review, we are interested in how nurses are leveraging their personal smartphones for personal and patient care–related purposes across a variety of health care settings.

This research aims to explore and synthesize the current literature regarding nurses' use of their personal digital devices in the workplace. The purpose of this review is to understand the current breadth of research to identify knowledge gaps, practice and policy implications, and future research opportunities. More specifically, this review seeks to gain an understanding of nurses' use of their personal digital technology within the workplace.

Methods

Overview

A scoping review was the chosen method to review the literature as it is well suited for burgeoning areas of research. In addition, scoping reviews aim to identify gaps in research while mapping key concepts and types of evidence [19]. Arksey and O'Malley's [20] methodological framework outlines five stages for conducting a scoping review, including identifying the research question; identifying relevant studies; study selection; charting the data; and collating, summarizing, and reporting the results. A critical step in this final stage involves identifying the research and evidence gaps that need to be addressed [20]. The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist was used to ensure that the scoping review is robust and includes all essential reporting items [21].

Step 1: Identify the Research Question

The research question we identified for this review is "How are nurses using their personal smartphones within the clinical workplace?"

Step 2: Identifying Relevant Studies

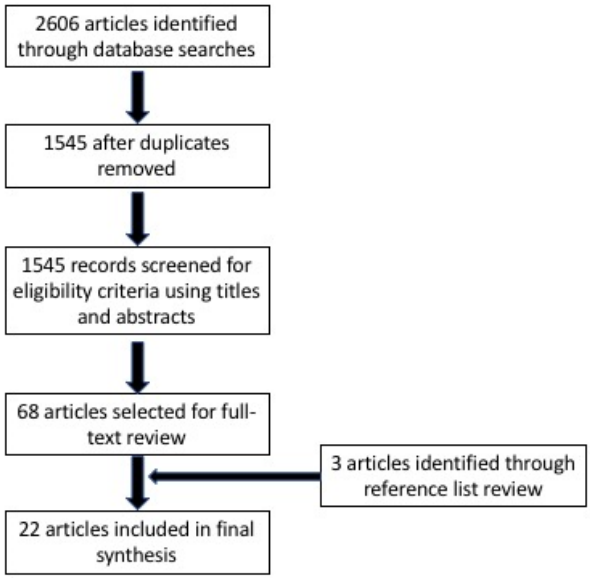
The following databases were used in the literature search process: CINAHL, PubMed, ProQuest Dissertations and Theses, Embase, MEDLINE, Nursing and Allied Health Database, Scopus, Web of Science, and Cochrane Reviews. Search terms used were: Nurs* AND (personal digital technology OR smartphone OR cellphone OR mobile phone OR cellular phone). Different search terms were trialed during the initial literature search process. The search terms selected were broad and were determined to capture all relevant studies that relate to the research question. See Multimedia Appendix 1 for the full syntax of search terms.

Articles included in the review had to meet the following inclusion criteria: a research focus on nurses' use of their own digital technology, research reported in English, and research published between January 2010 and January 2020 in order to capture data that aligned with the proliferation of mobile phone ownership. Studies focusing on a specific device or online app use were included only if the motivation for use was nurse driven. Studies were excluded from this review if the workplace organization provided the technology or if the technology was implemented for research purposes. Studies that focused on infection control of personal digital technology were excluded. Studies were also excluded if they focused on personal digital technology use and nursing students or nursing education, as this review was aimed at understanding practicing nurses' use of digital health technologies within clinical settings.

Step 3: Study Selection

A total of 2606 studies were identified using the search criteria above. Through consensus agreement between two reviewers, 22 studies met the inclusion and exclusion criteria and were included for the scoping review. See Figure 1 for a flow diagram of the article selection process.

Figure 1. Flow diagram for article selection process.



Step 4: Charting the Data

Two databases were set up in Step 4. One database, Mendeley, was used for sharing of articles between authors, including sharing to achieve consensus review. The other database was a matrix where all the data were collected. A spreadsheet was created to organize relevant information from the included articles. Data collected from each article included author names, year of publication, study design, data collection method, target population, study location, and research question.

Step 5: Collating, Summarizing, and Reporting the Results

During the final stage of collating, summarizing, and reporting the results, the framework approach described by Ritchie and Spencer [22] and used by Arksey and O'Malley [20] was used in this review. This technique involves synthesizing and interpreting the data by sifting, sorting, and charting the information based on the key themes and issues presented in the literature [22]. The data were charted and thematically analyzed. Data analysis was conducted by two researchers and analytical discrepancies were resolved through discussion until consensus was achieved.

Results

Overview

The 22 studies included in this research (see Tables 1-3 [23-44]) were conducted across a diverse range of countries and published between 2013 and 2019, with most (17/22, 77%) published in 2016 or later. Out of the 22 included studies, 16 (73%) used quantitative (ie, cross-sectional survey) research designs [23-38], 4 (18%) were qualitative [39-42], and 2 (9%) used a mixed methods design [43,44].

Table 1. Included quantitative studies.

Study design and author (year)	Data collection method	Target population	Study location	Research questions	Negative implications of personal digital technology	Purpose for digital technology use
Alameddine et al (2019) [23]	Survey, cross-sectional	97 emergency department (ED) providers, including ED faculty members, attending physicians, medical students, residents, and nurses (33% nurses)	Academic health center with the highest volume of patient visits in Lebanon	1. What are the frequency and patterns of smart device use among health care providers in the ED of a large academic health center in Lebanon?	• Distraction • Patient perception	• Information seeking for clinical use • Communication for clinical purposes • Personal communication purposes • Other personal uses
Bautista (2019) [24]	Pen and paper survey	517 staff nurses	19 tertiary-level general hospitals in Metro Manila, Philippines	1. Using a ranking system, how do Filipino nurses use their smartphones for work purposes? 2. What are the differences in nurses' smartphone use for work purposes based on demographic and organizational factors?	N/A[a]	• Communication for clinical purposes • Information seeking for clinical use
De Benedictis et al (2019) [25]	Survey	125 nurses and 66 physicians	An Italian university hospital in Rome, Italy	1. In what way is WhatsApp used in hospital settings by physicians and nurses with patients and between colleagues? 2. Which are the main perceived benefits and threats concerning the use of WhatsApp in a hospital setting by physicians and nurses? 3. Which are the determinants (individual and/or organizational) of the use of WhatsApp in a hospital setting? 4. Is there an interplay between individual and organizational determinants?	• Perceived risks (privacy and confidentiality) • Lack of organization regulations • Distraction	• Communication for clinical purposes • Personal communication purposes
Di Muzio et al (2019) [26]	Survey	193 nurses	Sapienza University Hospital, Rome, Italy	1. What is the validity and reliability of the Nurses' Use of PCDs[b] Questionnaire in the Italian hospitals?	• Increase error • Negative impact on performance • Distraction	• Personal communication purposes • Information seeking for clinical use • Other personal uses
Flynn et al (2018) [27]	Survey	735 acute care nurses, point of care and not point of care	Six acute care medical-surgical facilities in an urban health care system in the Southern United States	1. What are the current rates of personal smartphone use by nurses in acute care settings? 2a. What are nurses' preferences regarding the use of smartphone functionality in acute care settings? 2b. Are there differences in use by age category or role? 3a. What are nurses' perceptions of the benefits and drawbacks of using smartphones in the acute care setting? 3b. Are there differences in perceptions by age category or role?	• May upset families • Distraction • Increase error	• Information seeking for clinical use • Communication for clinical purposes • Personal communication purposes • Other personal uses

Study design and author (year)	Data collection method	Target population	Study location	Research questions	Negative implications of personal digital technology	Purpose for digital technology use
Garner et al (2017) [28]	Survey	97 acute care nurses and physicians (82.5% were nurses)	340+-bed tertiary facility in Bengaluru, India	1. What are the smartphone access and use, including future opportunities for mHealth[c] and potential ethical implications, among health care professionals practicing at a health care facility in Bengaluru, India?	• Confidentiality • Misuse of health information • Patient anxiety • Cybercrime • Limited Wi-Fi access	• Communication for clinical purposes • Personal communication purposes • Other personal uses
Grabowsky (2015) [29]	Survey	59 advanced practice nurses (APNs)	Alabama, United States; physician's office, outpatient clinics, hospitals, academic health centers, employee health clinic, hospice, nurse practitioner–owned practice, Veterans Administration nursing home, urgent care, health department, and dialysis unit	1. What types of clinical questions are answered using smartphones? 2. Are there barriers to information seeking with smartphones? 3. What phone apps and online resources do APNs find most useful in clinical situations? 4. How do APNs view their current online searching skills? 5. What is the level of interest in receiving training in online searching and what type of training is preferred? 6. Is the use of smartphones to answer clinical questions related to gender, level of education, population of practice area, practice type, or years approved to practice as an APN?	• Lack of internet access	• Information seeking for clinical use
Hranchook et al (2018) [30]	Survey	258 certified registered nurse anesthetists (CRNAs)	Michigan Association of Nurse Anesthetists, United States	1. What are the clinical and nonclinical uses of mobile computing devices among Michigan CRNAs? 2. What are the experiences of Michigan CRNAs with regard to the impact of using these devices on patient care?	• Distraction • Risk to patient • Performance decline • Policy	• Information seeking for clinical use • Communication for clinical purposes • Personal communication purposes • Other personal uses
Mayer et al (2019) [31]	Survey	1293 nurses across a range of settings, including hospital care, primary care, social health care, prehospital care, management, teaching and research, and private practice	Nursing Association of Barcelona	1. Are nurses using health apps professionally and what types of apps are they using? 2. Among nurses, is there a need for training in the use of health apps? 3. What are nurses' perceptions of health professional apps? 4. Is there a need for a certification process for health apps and what type of institution or organization should review and validate these apps for professional use?	• Concern about information quality in health apps	• Information seeking for clinical use
McBride et al (2015) [32]	Survey	825 acute care hospital registered nurses (RNs)	Members of Academy of Medical Surgical Nurses, United States	1. What is the frequency of non-work-related use of personal mobile phones and other personal communication devices among hospital RNs?	N/A	• Personal communication purposes • Other personal uses
	Survey					

Study design and author (year)	Data collection method	Target population	Study location	Research questions	Negative implications of personal digital technology	Purpose for digital technology use
McBride and LeVasseur (2017) [33]		1268 nurses (staff, charge, advanced practice, managers, faculty, and executive)	Members of the Academy of Medical Surgical Nurses and the Society of Pediatric Nurses, United States	1. How do RNs working on inpatient units use their PCDs at work (excluding lunch and breaks) and what are their opinions about how PCD use impacted their work and the work of their colleagues?	• Distraction • Negative impact on performance	• Information seeking for clinical use • Communication for clinical purposes • Personal communication purposes • Other personal uses
Mobasheri et al (2015) [34]	Survey	564 acute care nurses and 287 doctors	Five individual hospital sites in London, United Kingdom	1. What are the ways that frontline staff are using smartphones, tablet devices, and mHealth apps in the clinical environment?	• Privacy	• Communication for clinical purposes
Moore and Jayewaedene (2014) [35]	Survey	82 acute care nurses and 334 doctors	40+ acute trusts in England	1. How do nurses and doctors use their smartphones at work, what do they use them for, and do they assess the risks associated with the apps they use?	• Uncomfortable using phone in front of patients	• Information seeking for clinical use
Piscotty et al (2016) [36]	Survey	140 nurses in Registered Nurse to Bachelor of Science in Nursing program (RNs already had worked as nurses)	Public school of nursing in Southeast Michigan, United States	1. What is the prevalence of social media use by nurses during work hours?	N/A	• Personal communication purposes • Other personal uses
Pucciarelli et al (2019) [37]	Survey	256 acute care nurses who worked in hospitals, outpatient facilities, or day surgeries	Seven hospitals in Central and Southern Italy	1. What are the work- and non-work-related activities performed by nurses using smartphones in the workplace? 2. What are the differences between smartphone use and nurses' age, gender, and working environment? 3.What are the positive or negative influences that smartphones have on nurses' performance during their health care activities?	• Distraction • Negative impact on performance • Inappropriate recording • Inappropriate vital signs measuring	• Information seeking for clinical use • Communication for clinical purposes • Personal communication purposes • Other personal uses
Stergiannis et al (2017) [38]	Survey	974 acute care medical and nursing staff (18.5% nursing assistants and 42.6% nurses, with the rest being doctors and junior doctors)	Six general hospitals in Athens, Greece	1. What is the clinical use of smartphones among medical and nursing staff in Greece?	• Unaware of apps that can be used to assist them in their daily clinical task • Internet access • Did not think smartphones were useful • Unsure about appropriate sites and apps • Distrustful of information • Lack of education on how to use phone	N/A

[a]N/A: not applicable; this information was not reported in the study.

[b]PCD: personal communication device.

[c]mHealth: mobile health.

Table 2. Included qualitative studies.

Study design and author (year)	Data collection method	Target population	Study location	Research questions	Negative implications of personal digital technology	Purpose for digital technology use
Bautista and Lin (2016) [40]	Semistructured interview	30 acute care staff, charge, and nurse managers	13 tertiary hospitals in the Philippines	1. How do the interactions of sociotechnical components (users, technology, and policy) affect staff nurses' use of personal mobile phones at work?	• Personal cost, as hospital does not provide phone • Distraction • Privacy gaps and punishments • Policy • Patient complaints • Infection control	• Information seeking for clinical use • Communication for clinical purposes
Bautista and Lin (2017) [39]	Semistructured interview	20 acute care staff nurses	Nine hospitals in the Philippines	1. How and why are mobile instant messaging apps used by Filipino nurses as part of their work? 2. What are the gratifications derived by nurses when using mobile instant messaging apps?	N/A[a]	• Communication for clinical purposes • Personal communication purposes
Chiang and Wang (2016) [41]	Semistructured interview	17 community nurses working for home care facilities	Two regional hospital-affiliated home care facilities and four community home care facilities in Southern Taiwan	1. What are nurses' experiences regarding the benefits and obstacles of using a smart mobile device app in home care?	• Perceived risks (privacy and confidentiality) • Lack of organization regulations and incentives • Disturbance to personal life, as messages received during nonworking hours	• Communication for clinical purposes
Park and Lee (2019) [42]	Semistructured focus group interviews	4 orthopedic scrub nurses	Operating room of a veterans' hospital, Korea	1. How are scrub nurses using a commercially available smartphone app to solve information needs in orthopedic surgery?	N/A	• Using an app to seek information for clinical use

[a]N/A: not applicable; this information was not reported in the study.

Table 3. Included mixed methods studies.

Study design and author (year)	Data collection method	Target population	Study location	Research questions	Negative implications of personal digital technology	Purpose for digital technology use
Giles-Smith et al (2017)[a] [43]	Survey, focus groups	94 acute care inpatient medical and surgical nurses	Community hospital and tertiary hospital in Winnipeg, Canada	1. What is the current usage of mobile devices and apps by nurses for direct patient care within the study sites? 2. What are the attitudes of nurses at these study sites toward the use of mobile devices and apps for direct patient care?	• Lack of wireless internet • Distraction • Unsure if allowed to use at work • Potential damage or loss of phone • Infection control • Concern of patient perception of phone • Professionalism	• Information seeking for clinical use
Planitz et al (2013) [44]	Survey, observation	299 acute care nurses	Various units and wards in a tertiary hospital located in Brisbane, Australia	1. What is the level of actual personal smartphone use by nurses that is occurring within the hospital? 2. What are the attitudes of nurses toward smartphone use at the hospital? 3. Do the benefits of smartphones outweigh the consequences of distractions and occasional misuse? 4. What are the factors influencing whether nurses used smartphones at work? 5. How are nurses using their smartphones at work?	• Disruptive	• Personal communication purposes • Other personal uses

[a]Two-part study: only part 1 relates to the research question.

Health Care Settings

In reviewing the included studies, most researchers focused on investigating personal smartphone use among hospital-based nurses (18/22, 82%) [23-28,30,32-35,37-40,42-44]. As well, 6 studies out of 22 (27%) included a multidisciplinary group that also included inquiry into physicians' use of personal smartphones in the clinical setting [23,25,28,34,35,38]. There was a single study that investigated registered nurses' personal technology use during work hours, whose sample population were enrolled in a Registered Nurse to Bachelor of Science in Nursing program, with 93% of the nurse participants currently working in acute care [36].

Out of 22 studies, 2 (9%) targeted nurses working in community-based health care settings [29,41]. Grabowsky [29] focused on advanced practice nurses who worked in an array of settings, including a variety of outpatient clinics. Chiang and Wang [41] investigated personal digital technology use among community nurses who worked for home care organizations. Out of 22 studies, 1 (5%) included nurses from a range of settings, both acute and nonacute, including hospital, primary care, social health care, prehospital care, management, education, and private practice settings [31].

Two main themes with associated subthemes were generated from the analyses. The first theme was *personal smartphone use for patient care*. This theme was divided into three subthemes: information seeking, communication, and mobile device functions. The second theme was *implications of personal smartphone use*, which was refined into five subthemes: smartphone use for personal reasons, distraction, patient perception, privacy and confidentiality, and organizational support and policy confusion. The details of the themes and subthemes are described narratively in the following sections.

Personal Smartphone Use for Patient Care

Information Seeking

Nurses used their smartphones to access information directed at patient care. The most common information sought was related to medications, including drug guides and drug references, in 14 of the 22 (64%) studies reviewed [23,26-31,33-35,37,38,40,43]. Nurses reported efficiency of access to medication information; it was easier to search medications by either the generic name or brand name using their smartphones, despite having access to a drug reference book [40]. By using a mobile device, nurses were able to quickly look up new medications in drug guides [43]. Out of 22 studies, 8 (36%) identified that nurses used apps downloaded to their devices to find information on medications; 4 (18%) did not report the name of the apps used [23,27,34,35]. Whether nurses located medication information though a general search engine app (eg, Google) or by a specific app was not disclosed. The other 4 (18%) studies identified the specific apps used by nurses

as Epocrates, Micromedex, Drug Index and Dosage, and Vademecum International [28,29,31,43].

Nurses also used their smartphones to access the internet for disease-related information [27,29,30,34,40,43]. Smartphones were used to collect procedural information, including various surgical procedures and anesthetic procedures [27,30,34,40,42]. Other than Park and Lee [42], researchers did not specify if nurses located this information specifically through online search engines or by specific apps they downloaded to their smartphone devices.

Nurses identified the need for information access regarding procedures and use of instruments within surgical settings [42]. The greatest barrier to information access included the inconvenience created by a lack of dedicated computers within surgical suites, thereby limiting nurses' ability to access information [42]. Surgical nurses used the app, BAND, on their personal devices in order to consult with each other and to share information, including pictures, videos of themselves describing instruments and procedures, YouTube videos, and other pertinent information [42]. Using the app, information was easy to access and readily available and was positively reviewed by the operating room nurses [42].

Nurses used their smartphones for information seeking for both patient education purposes and for their own educational needs. Out of 22 studies, 4 (18%) reported that nurses used their smartphones for patient education [27,29,33,40]. Flynn et al [27] found that more than 75% of the nurses preferred the use of their smartphones to access information for patient education, and this was more commonly noted among nurses between the ages of 18 and 30 years compared to nurses older than 50 years of age. Furthermore, Flynn et al [27] discussed how nurses used their own smartphones to respond to patients' and families' requests for information on a variety of topics, such as medical procedures, medications, contact information for providers, and directions to local venues.

Personal smartphone use provided nurses with ease of access to information. For example, when asked by a patient about lab values, the nurse used their phone to search Google for the answer in the moment [40]. McBride and Levasseur [33] reported that nurses also used their personal smartphones to access patient handouts and teaching materials. Nurses accessed continuing education and professional development opportunities through their smartphones [30,33], and Mobasheri et al [34] reported that nurses used medical apps to enhance their clinical knowledge and skills.

Other uses of smartphones by nurses included the use of online evidenced-based guidelines to support patient care [26,30], to access information related to patients' prognoses [29], to support nurses' understanding of disease pathophysiology and physiology [30], and to access clinical decision support tools [35]. Multiple smartphone apps that were reported being used by nurses for information seeking included WebMD [28,40], Medscape [23,28,40], Google [28,40,43], UpToDate [23], Infermera virtual [31], and 061 CatSalut Respon [31]. Garner et al [28] reported nurses' use of PubMed, YouTube, Medical in Nursing and Oncology App, Google Scholar, Q Calc, Doc Plus, Praco, Radiopedia, and Wikipedia as apps and programs on their smartphones. Giles-Smith et al [43] identified Lexicomp, Medscape, and iTriage as additional health-related apps on nurses' smartphones for information seeking and patient care purposes. Nurses, particularly those with less than 5 years of clinical experience, used e-books on their personal smartphones for information seeking [24].

Communication

Nurses used their personal smartphones to communicate with members of the health care team. Out of 22 studies, 11 (50%) found that nurses used their smartphones while at work to connect with health care team members by calling and/or text messaging for patient care purposes [24-28,30,33,34,37,39,40]. Nurses reported improved communication between team members, including physicians, and other allied health care providers and claimed that their personal device use improved efficiency in communication and facilitated immediate contact with colleagues [27,39,40]. Voice calling and text messaging with nurses and physicians was the most commonly cited reason for personal smartphone use at work [24].

Bautista and Lin [40] reported that nurses were able to contact doctors immediately using the Viber app and also received patient care orders via their smartphones. Nurses reported an enhanced efficiency in their workflow because they did not need to "track down" the attending physician; instead, they used their smartphones to contact them immediately [39]. Another efficiency-related technique was the development of a directory of personal contacts, whereby nurses saved the contact information (eg, phone numbers) of physicians, supervisors, and colleagues across diverse organizational departments to their smartphones so they were readily available when required [27]. Specific apps used for communication among the health care team included Viber, Facebook Messenger, iMessage, WhatsApp, and Line [25,28,37,39]. Group chats were also mentioned, but a particular app was not reported for group chat purposes [39].

Nurses also used their smartphones to communicate with patients [41]. Chiang and Wang [41] assessed nurses' experiences using the Line app to communicate with home care patients and families. Home care patients or their family members would contact the nurse regarding their health care via the text option on the Line app [41]. Chiang and Wang [41] found that the picture and video function on the app was used as an assessment tool; home care patients would take pictures and videos of signs and symptoms related to their health condition and send them to the nurse using the app. Nurses found the Line app beneficial because of the asynchronous nature of communication patterns, which afforded them time to problem solve patient care issues, and the unlimited texting function increased opportunities for patient-provider communication [41]. However, nurses also reported a situation of being chronically "on call," where patient text messages were perceived as intrusive outside of their working hours [41]. In another study, which investigated the use of WhatsApp, researchers found that nurses rarely communicated with patients, whereas in the same study, physicians were more likely to communicate with patients using WhatsApp [25].

Mobile Device Functions

Three smartphone functions were identified as useful: the photo and video function [24,27,34,39,40], the calculator [26,27,31,33-35,37,38], and the flashlight function [27,44]. Nurses used the photo and video function on their smartphones to support communication with their colleagues. For example, nurses took pictures of radiology images, patient wound sites, electrocardiograms, procedural equipment, skin test results, and biological samples, such as sputum [27,40], to facilitate communication with other care providers who requested greater details of the patient condition beyond verbal reports [27].

Flynn et al [27] reported that over 75% of the nurses in their study valued the calculator function on their smartphones. The calculator was used to compute nursing and medical formulas and to calculate dose and scale [26,31,37]. Other than reported use, no further details were provided regarding the flashlight function.

Implications of Personal Smartphone Use

Smartphone Use for Personal Reasons

In 12 out of 22 (55%) studies, nurses used their smartphones within the workplace setting for nonwork-related reasons [23,25-28,30,32,33,36,37,39,44]. Nurses' most commonly reported smartphone use included texting and messaging family and friends [25,27,28,30,32,33,36,37,39,44]. The second-most common personal smartphone use was to check or post information on their personal social media sites [23,26,28,30,32,33,36,37,44]. Other common uses were for telephone calling or checking for missed calls [27,28,30,36,44] and for entertainment such as playing games or shopping [26,32,33,44]. Personal use was reported outside of scheduled breaks [32,33]. Uses that were only mentioned once throughout the literature included internet access and surfing [30], catharsis and conveying grievances about their work with other nurses [39], the use of the global positioning system to verify their children's location [27], and for personal safety (ie, walking to vehicles at night) [27].

Distraction

Distraction from patient care was identified by nurses as a concern regarding smartphone use within the workplace [23,25-27,30,33,37,40,43]. Nurses self-reported instances of their own distraction from patient care activities resulting from the use of their smartphones but also reported observed instances of inattentiveness among their nursing colleagues [23,26,30,33,37]. Pucciarelli et al [37] reported that 42% of nurses felt that smartphones were a distraction. For example, a nurse was distracted from inserting an intravenous line by the ringing sound of an incoming call on their smartphones [40], and 12.5% of nurses had witnessed a coworker make a medical error that they attributed to perceived distraction related to their smartphone use [33]. Alameddine et al [23] similarly found that 55% of nurses had witnessed distraction they attributed to smartphone use. Di Muzio et al [26] found that almost 62% of nurses thought that the use of personal devices could increase the risk of errors.

Patient Perception

Similarly, nurses expressed their lack of knowledge about patients' perceptions of their smartphone use and the potential for patient complaints [27,28,35,40,43]. Nurses were concerned about patients' possible negative feelings and thoughts about mobile devices being used at the beside [23,43].

Privacy and Confidentiality

Nurses identified privacy and confidentiality of patient health information as a concern associated with the use of their personal devices within the workplace [25,28,34,40,41]. Mobasheri et al [34] found that 3.6% of nurses believed there was patient-related clinical information retained on their personal smartphones. Similarly, home care nurses discussed their concerns regarding their use of the Line app, reporting that "others could easily see patients' recent whereabouts or sensitive personal information on the software platform" [41].

Lack of Organizational Support and Policy Confusion

Lack of organizational support and policy confusion was expressed by nurses [30,40,41,43]. Despite organizational policy that prohibited personal smartphone use within their clinical setting, most nurses considered their smartphones as helpful technology in their workplace. Some nurses were aware of policies restricting personal smartphone use, while others stated they did not know what the existing policies for personal smartphone use entailed. Hranchook et al [30] reported that 47.2% of the nurse participants in their study knew that their organization had a personal device use policy, 22.6% said their institution did not, while 30.2% were unsure. When asked if medical administrators should develop a code of conduct for smartphone use to minimize unnecessary distraction, 44% of participants agreed, while 51% disagreed [23]. Nurses also reported barriers to Wi-Fi and internet access issues within their organizations that impeded their access to online resources [28,29,38,43].

Discussion

Principal Findings

A scoping review of 22 studies was conducted to examine the current breadth and range of research on nurses' use of personal smartphones in the workplace; to our knowledge, this is the first review of its kind. Based on the studies included in this review, nurses have reportedly been using their personal smartphones within their workplace for personal and patient care purposes since 2013.

From the existing evidence, we know that nurses used their personal smartphone devices to gather patient care information and to communicate within the health care team. In support of evidence-informed patient care, nurses used online apps and programs to locate information relating to patient-prescribed medications, clinical procedures, diagnoses, laboratory tests, and more. The information-seeking and consultation behaviors align with evidence-informed nursing practice [18,45]. Nurses require access to specialized knowledge and comprehensive clinical information to inform their clinical decision making. However, the significant demands on nurses' time impedes their

ability to engage with research and health care resources required for evidence-informed practice [46]. The use of their personal smartphone devices provided nurses with rapid and easy access to online health information. It is possible that nurses are using their personal smartphones as a way to efficiently access needed resources and support their patient care information needs.

Despite the high level of smartphone and app use, there was minimal reporting on nurses' attention to the quality and accuracy of the information garnered from the apps used. The commercial profitability of health-related apps tend to take priority over ensuring that apps are critically appraised for information accuracy [47]. However, when asked, nurses expressed that they want health apps to be certified by a health or professional institution, exemplifying their concern about information quality [31]. Nurses who conducted a risk assessment on the apps that they used assessed the trustworthiness of the source, sought guidance from professional bodies on technology and app use, and were vigilant about patient information privacy [35].

Similarly, nurses may be resorting to their own personal digital devices to compensate for the lack of support provided by their organizations [42]. As reported elsewhere, nurses who have access to resources from a medical library, have internet access at work, and have the opportunity to work with computer technologies are more likely to engage in evidence-informed nursing practice [48]. Organizations that strive to provide the highest-quality patient care are also accountable to providing the resources that make this achievable. The presupposition of health information technology within the clinical setting highlights the need for organizations to provide continuing education to all health care providers regarding risk assessments of health information technologies (eg, devices and apps) for responsible use and to support evidence-informed practices.

Effective communication among the health care team is essential for quality patient care and effective teamwork [49]. In fact, the most common reason for errors resulting in patient harm result from communication failures [50]. In our review, nurses used text messaging, telephone, and picture and video functions on their smartphones to communicate with other members of the health care team. Furthermore, nurses preferred the use of their own smartphones for efficient and immediate communication among health care team members. Medical students have also reported using personal smartphones as an efficient means of communication and for coordination of the clinical team [51]. It may be that organizationally based communication systems do not provide satisfactory options to meet the precipitous communication needs of health care teams working within a digital health context [52]. Similarly, nurses' use of their personal devices to communicate among members of the health care team may demonstrate a technological "work-around" meant to reconcile health system demands for cost-efficiency with efforts to provide quality patient care [53].

Despite nurses' desires to use their personal smartphones in the workplace, they were also mindful of concerns that included personal and/or colleague distraction, and they acknowledged challenges to information privacy and information security related to possible retention of patient data on their smartphones. Nurses' use of their personal devices to share practice-related information among health care teams may characterize their "work-around" of organizational health information systems that have not effectively considered nurses' workflows, their need for accessible and up-to-date health care information, nor communication practices among multidisciplinary health care teams [52,54]. Nurses are not the only health professionals experiencing these tensions. Physicians report that barriers to using their smartphones in practice include organizational policies against smartphone use and concerns about using them during patient consultations [9]. The tension between quality health practices and pressure to find system efficiencies may have also contributed to nurses' use of their personal smartphones within the workplace [54]. The reported lack of awareness of organizational policies related to personal smartphone use and nurses' lack of knowledge regarding patient perception of their smartphone use for patient care constitute two important areas for future research and health professional education [55].

Implications

Implications for Research

This scoping review highlights significant gaps in research regarding nurses' use of personal smartphones in the workplace. To date, research has focused on acute care nurses, with little investigation of nurses' device use within other practice settings. Research is needed to understand personal smartphone use across a more diverse range of health care settings, such as home care, long-term care, and public health. A greater understanding of other health care settings, such as long-term care, will provide invaluable information, especially due to the aging population in Canada and the growing demand on care providers for the elderly. COVID-19 exposed profound gaps in long-term care in Canada, including access to practice standards and education. The importance of public health care workers was also emphasized during the COVID-19 pandemic. Research can help to inform how smartphones can be leveraged to support nurses, especially during resource and staffing shortages. This will also help to generate a greater understanding of the online resources (eg, online information sites and apps) that support nurses' practices. Research is also needed to inform smartphone practice strategies and policies that support efficient team communication and that are accountable to information, privacy, and security concerns.

Implications for Practice

Nurses are using their personal devices to access clinical information and to contribute to clinical decision making. If apps are being used and patient information is entered, it is important to know where the information is being stored, who has access to it, and who owns the data. Nurses' use of their personal digital devices for sharing patient-related information within the health care team carries a potential risk for inappropriate disclosure of personal health information. While unapologetic about using their personal devices, nurses instead advocated for information-secure messaging apps that would ensure the safety and security of people's personal health information [34].

Implications for Policy

Evidence that over 75% of nurses and physicians within acute care settings use their personal smartphones to support their patient care practices is indicative of a clear gap between policy and practice [27,38]. If nurses, as well as other health care professionals, use their smartphones as an efficient way to retrieve and share information, it would be incumbent on organizations to reassess their policies regarding personal phone use and to work with technology developers and care providers to nurture their innovative thinking and insight into effective, responsive, and responsible technology-enabled solutions.

Limitations

There were several limitations when completing this scoping review. As 16 of the 22 studies were quantitative and used cross-sectional surveys to gather data, there were limited descriptions and examples available to provide an in-depth understanding of smartphone and mobile phone use. In addition, due to the lack of diversity in the target population, there is limited evidence across the nursing continuum. Lastly, studies included were only in English; studies in other languages were not included in these findings.

Conclusions

This scoping review provides insight into nurses' use of their personal smartphones within the workplace setting. The review highlights knowledge gaps regarding nurses' personal device use and the safety and privacy of personal health information, patient care outcomes, communication practices among health care teams, and insight into app and information accuracy and reliability, thus creating significant opportunities for future research regarding nurses' use of personal digital technology in the workplace. Organizational policies that limit or prohibit smartphone use may be shortsighted; insightful leadership would leverage the knowledge of nurses with other health care providers to collaboratively develop strategies that enable efficient, respectful, and ethical use of communication technology for effective patient care practices [56].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full syntax of search terms for the literature search process.
[PDF File (Adobe PDF File), 73 KB - mhealth_v8i11e18774_app1.pdf]

References

1. Number of smartphone users from 2016 to 2021. Statista. 2019. URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [accessed 2019-07-03]
2. Mosa ASM, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012 Jul 10;12(1):67 [FREE Full text] [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]
3. Kuss D, Harkin L, Kanjo E, Billieux J. Problematic smartphone use: Investigating contemporary experiences using a convergent design. Int J Environ Res Public Health 2018 Jan 16;15(1):142 [FREE Full text] [doi: 10.3390/ijerph15010142] [Medline: 29337883]
4. Ozdalga E, Ozdalga A, Ahuja N. The smartphone in medicine: A review of current and potential use among physicians and students. J Med Internet Res 2012 Sep 27;14(5):e128 [FREE Full text] [doi: 10.2196/jmir.1994] [Medline: 23017375]
5. Internet usage frequency in Canada as of January 2018. Statista. 2018. URL: https://www.statista.com/statistics/686835/canada-internet-usage-frequency/%0Ale [accessed 2018-07-03]
6. Charani E, Castro-Sánchez E, Moore LS, Holmes A. Do smartphone applications in healthcare require a governance and legal framework? It depends on the application!. BMC Med 2014 Feb 14;12:29 [FREE Full text] [doi: 10.1186/1741-7015-12-29] [Medline: 24524344]
7. Fölster S. Viral mHealth. Glob Health Action 2017 Jun;10(sup3):1336006 [FREE Full text] [doi: 10.1080/16549716.2017.1336006] [Medline: 28838304]
8. Canadian Institute for Health Information. Better Information for Improved Health: A Vision for Health System Use of Data in Canada. Ottawa, ON: Canadian Institute for Health Information; 2013 Jun. URL: https://www.cihi.ca/sites/default/files/document/hsu_vision_report_en.pdf [accessed 2019-04-13]
9. Byambasuren O, Beller E, Glasziou P. Current knowledge and adoption of mobile health apps among Australian general practitioners: Survey study. JMIR Mhealth Uhealth 2019 Jun 03;7(6):e13199 [FREE Full text] [doi: 10.2196/13199] [Medline: 31199343]
10. Jedwab RM, Chalmers C, Dobroff N, Redley B. Measuring nursing benefits of an electronic medical record system: A scoping review. Collegian 2019 Oct;26(5):562-582. [doi: 10.1016/j.colegn.2019.01.003]
11. Akhu-Zaheya L, Al-Maaitah R, Bany Hani S. Quality of nursing documentation: Paper-based health records versus electronic-based health records. J Clin Nurs 2018 Feb;27(3-4):e578-e589. [doi: 10.1111/jocn.14097] [Medline: 28981172]
12. Top M, Gider O. Nurses' views on electronic medical records (EMR) in Turkey: An analysis according to use, quality and user satisfaction. J Med Syst 2012 Jun;36(3):1979-1988. [doi: 10.1007/s10916-011-9657-6] [Medline: 21302133]

13. Mobasheri MH, King D, Johnston M, Gautama S, Purkayastha S, Darzi A. The ownership and clinical use of smartphones by doctors and nurses in the UK: A multicentre survey study. BMJ Innov 2015 Oct 07;1(4):174-181. [doi: 10.1136/bmjinnov-2015-000062]
14. McBride DL, LeVasseur SA. Personal communication device use by nurses providing in-patient care: Survey of prevalence, patterns, and distraction potential. JMIR Hum Factors 2017 Apr 13;4(2):e10 [FREE Full text] [doi: 10.2196/humanfactors.5110] [Medline: 28408359]
15. Moore S, Anderson J, Cox S. Pros and cons of using apps in clinical practice. Nurs Manag (Harrow) 2012 Oct;19(6):14-17. [doi: 10.7748/nm2012.10.19.6.14.c9316] [Medline: 23252086]
16. Naming the coronavirus disease (COVID-19) and the virus that causes it. World Health Organization. Geneva, Switzerland: World Health Organization; 2020. URL: https://tinyurl.com/y22lpyod [accessed 2020-07-02]
17. Canadian Nurses Association. Evidence-Informed Decision-Making and Nursing Practice. Ottawa, ON: Canadian Nurses Association; 2010. URL: https://www.cna-aiic.ca/-/media/nurseone/page-content/pdf-en/evidence-informed-decision-making-and-nursing-practice.pdf [accessed 2019-04-20]
18. Canadian Nurses Association. Evidence-Informed Decision-Making and Nursing Practice. Ottawa, ON: Canadian Nurses Association; 2018. URL: https://www.cna-aiic.ca/-/media/cna/page-content/pdf-en/evidence-informed-decision-making-and-nursing-practice-position-statement_dec-2018.pdf [accessed 2020-11-09]
19. Colquhoun HL, Levac D, O'Brien KK, Straus S, Tricco AC, Perrier L, et al. Scoping reviews: Time for clarity in definition, methods, and reporting. J Clin Epidemiol 2014 Dec;67(12):1291-1294. [doi: 10.1016/j.jclinepi.2014.03.013] [Medline: 25034198]
20. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
21. 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]
22. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG, editors. Analyzing Qualitative Data. Abingdon, UK: Routledge; 1994.
23. Alameddine M, Soueidan H, Makki M, Tamim H, Hitti E. The use of smart devices by care providers in emergency departments: Cross-sectional survey design. JMIR Mhealth Uhealth 2019 May 05;7(6):e13614 [FREE Full text] [doi: 10.2196/13614] [Medline: 31199328]
24. Bautista JR. Filipino nurses' use of smartphones in clinical settings. Comput Inform Nurs 2019 Feb;37(2):80-89. [doi: 10.1097/CIN.0000000000000482] [Medline: 30299305]
25. De Benedictis A, Lettieri E, Masella C, Gastaldi L, Macchini G, Santu C, et al. WhatsApp in hospital? An empirical investigation of individual and organizational determinants to use. PLoS One 2019;14(1):e0209873 [FREE Full text] [doi: 10.1371/journal.pone.0209873] [Medline: 30633754]
26. Di Muzio M, Napoli C, Fiorinelli M, Battista Orsi G. Nurses distracted by smartphone use during clinical practice: Italian validation of the "Nurses' use of Personal Communication Devices (PCD) questionnaire". Appl Nurs Res 2019 Oct;49:23-28. [doi: 10.1016/j.apnr.2019.07.008] [Medline: 31495415]
27. Flynn GAH, Polivka B, Behr JH. Smartphone use by nurses in acute care settings. Comput Inform Nurs 2018 Mar;36(3):120-126. [doi: 10.1097/CIN.0000000000000400] [Medline: 29251648]
28. Garner SL, Sudia T, Rachaprolu S. Smart phone accessibility and mHealth use in a limited resource setting. Int J Nurs Pract 2018 Feb;24(1):e12609. [doi: 10.1111/ijn.12609] [Medline: 29159919]
29. Grabowsky A. Smartphone use to answer clinical questions: A descriptive study of APNs. Med Ref Serv Q 2015;34(2):135-148. [doi: 10.1080/02763869.2015.1019320] [Medline: 25927507]
30. Hranchook A, Penprase BB, Piscotty RJ. Mobile computing devices in the perioperative environment: A survey exploring uses and experiences among certified registered nurse anesthetists. AANA J 2018 Dec;86(6):471-478. [Medline: 31584421]
31. Mayer M, Rodríguez Blanco O, Torrejon A. Use of Health Apps by Nurses for Professional Purposes: Web-Based Survey Study. JMIR Mhealth Uhealth 2019 Nov 1;7(11):e15195-e15E50. [doi: 10.2196/15195] [Medline: 22861175]
32. McBride DL, LeVasseur SA, Li D. Non-work-related use of personal mobile phones by hospital registered nurses. JMIR Mhealth Uhealth 2015 Jan 13;3(1):e3 [FREE Full text] [doi: 10.2196/mhealth.4001] [Medline: 25586982]
33. McBride DL, LeVasseur SA. Personal communication device use by nurses providing in-patient care: Survey of prevalence, patterns, and distraction potential. JMIR Hum Factors 2017 Apr 13;4(2):e10 [FREE Full text] [doi: 10.2196/humanfactors.5110] [Medline: 28408359]
34. Mobasheri MH, King D, Johnston M, Gautama S, Purkayastha S, Darzi A. The ownership and clinical use of smartphones by doctors and nurses in the UK: A multicentre survey study. BMJ Innov 2015 Oct 07;1(4):174-181. [doi: 10.1136/bmjinnov-2015-000062]
35. Moore S, Jayewardene D. The use of smartphones in clinical practice. Nurs Manag (Harrow) 2014 Jul;21(4):18-22. [doi: 10.7748/nm.21.4.18.e1225] [Medline: 24967805]
36. Piscotty R, Martindell E, Karim M. Nurses' self-reported use of social media and mobile devices in the work setting. J Nurs Infom 2016;20(1):9.

37. Pucciarelli G, Simeone S, Virgolesi M, Madonna G, Proietti MG, Rocco G, et al. Nursing-related smartphone activities in the Italian nursing population. Comput Inform Nurs 2019;37(1):29-38. [doi: 10.1097/cin.0000000000000474]
38. Stergiannis P, Intas G, Toulia G, Tsolakoglou I, Kostagiolas P, Christodoulou E, et al. Clinical use of smartphones among medical and nursing staff in Greece. Comput Inform Nurs 2017;35(9):483-488. [doi: 10.1097/cin.0000000000000347]
39. Bautista JR, Lin TT. Nurses' use of mobile instant messaging applications: A uses and gratifications perspective. Int J Nurs Pract 2017 Oct;23(5):e12577. [doi: 10.1111/ijn.12577] [Medline: 28752519]
40. Bautista JR, Lin TT. Sociotechnical analysis of nurses' use of personal mobile phones at work. Int J Med Inform 2016 Nov;95:71-80. [doi: 10.1016/j.ijmedinf.2016.09.002] [Medline: 27697234]
41. Chiang K, Wang H. Nurses' experiences of using a smart mobile device application to assist home care for patients with chronic disease: A qualitative study. J Clin Nurs 2016 Jul;25(13-14):2008-2017. [doi: 10.1111/jocn.13231] [Medline: 27136280]
42. Park H, Lee S. International nursing: Use of a commercially available smartphone application to solve information needs of orthopedic scrub nurses. Nurs Adm Q 2019;43(4):337-350. [doi: 10.1097/NAQ.0000000000000366] [Medline: 31479055]
43. Giles-Smith L, Spencer A, Shaw C, Porter C, Lobchuk M. A study of the impact of an educational intervention on nurse attitudes and behaviours toward mobile device use in hospital settings. J Can Health Libr Assoc 2017 Apr 06;38(1):12-29 [FREE Full text] [doi: 10.5596/c17-003]
44. Planitz B, Sanderson P, Kipps T, Driver C. Nurses' self-reported smartphone use during clinical care. In: Proceedings of the 57th Human Factors and Ergonomics Society Annual Meeting. 2013 Sep 30 Presented at: 57th Human Factors and Ergonomics Society Annual Meeting; September 30-October 4, 2013; San Diego, CA p. 738-742. [doi: 10.1177/1541931213571161]
45. Snyder-Halpern R, Corcoran-Perry S, Narayan S. Developing clinical practice environments supporting the knowledge work of nurses. Comput Nurs 2001;19(1):17-23; quiz 24. [Medline: 11210450]
46. Belita E, Yost J, Squires JE, Ganann R, Burnett T, Dobbins M. Measures assessing attributes of evidence-informed decision-making (EIDM) competence among nurses: A systematic review protocol. Syst Rev 2018 Nov 03;7(1):181 [FREE Full text] [doi: 10.1186/s13643-018-0849-8] [Medline: 30390711]
47. Moore S, Anderson J, Cox S. Pros and cons of using apps in clinical practice. Nurs Manag (Harrow) 2012 Oct;19(6):14-17. [doi: 10.7748/nm2012.10.19.6.14.c9316] [Medline: 23252086]
48. Eizenberg M. Implementation of evidence-based nursing practice: Nurses' personal and professional factors? J Adv Nurs 2011 Jan;67(1):33-42. [doi: 10.1111/j.1365-2648.2010.05488.x] [Medline: 20969620]
49. Garrett JH. Effective perioperative communication to enhance patient care. AORN J 2016 Aug;104(2):111-120. [doi: 10.1016/j.aorn.2016.06.001] [Medline: 27472971]
50. Leonard M, Graham S, Bonacum D. The human factor: The critical importance of effective teamwork and communication in providing safe care. Qual Saf Health Care 2004 Oct;13 Suppl 1:i85-i90 [FREE Full text] [doi: 10.1136/qhc.13.suppl_1.i85] [Medline: 15465961]
51. Shenouda JEA, Davies BS, Haq I. The role of the smartphone in the transition from medical student to foundation trainee: A qualitative interview and focus group study. BMC Med Educ 2018 Jul 31;18(1):175 [FREE Full text] [doi: 10.1186/s12909-018-1279-y] [Medline: 30064424]
52. Evans RS. Electronic health records: Then, now, and in the future. Yearb Med Inform 2018 Mar 06;25(S 01):S48-S61. [doi: 10.15265/iys-2016-s006]
53. Kieft RA, de Brouwer BB, Francke AL, Delnoij DM. How nurses and their work environment affect patient experiences of the quality of care: A qualitative study. BMC Health Serv Res 2014 Jun 13;14:249 [FREE Full text] [doi: 10.1186/1472-6963-14-249] [Medline: 24923663]
54. Salmond SW, Echevarria M. Healthcare transformation and changing roles for nursing. Orthop Nurs 2017;36(1):12-25. [doi: 10.1097/nor.0000000000000308]
55. Powell-Cope G, Nelson A, Patterson E. Patient care technology safety. Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville, MD: Agency for Healthcare Research and Quality, US Department of Health and Human Services; 2008 Apr. URL: https://archive.ahrq.gov/professionals/clinicians-providers/resources/nursing/resources/nurseshdbk/nurseshdbk.pdf [accessed 2020-11-09]
56. Brandt J, Katsma D, Crayton D, Pingenot A. Calling in at work. Nurs Manag 2016;47(7):20-27. [doi: 10.1097/01.numa.0000484474.76112.a3]

Abbreviations

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
PRN: *pro re nata*

Edited by G Eysenbach, L Buis; submitted 17.03.20; peer-reviewed by G Strudwick, J Ayre; comments to author 12.06.20; revised version received 28.07.20; accepted 19.10.20; published 26.11.20.

Please cite as:
de Jong A, Donelle L, Kerr M
Nurses' Use of Personal Smartphone Technology in the Workplace: Scoping Review
JMIR Mhealth Uhealth 2020;8(11):e18774
URL: https://mhealth.jmir.org/2020/11/e18774
doi:10.2196/18774
PMID:33242012

Original Paper

Understanding User Behavior Through the Use of Unsupervised Anomaly Detection: Proof of Concept Using Internet of Things Smart Home Thermostat Data for Improving Public Health Surveillance

Niloofar Jalali[1], MSc, PhD; Kirti Sundar Sahu[1], MPH, MPT; Arlene Oetomo[1], BSc; Plinio Pelegrini Morita[1,2,3,4], MSc, PEng, PhD

[1]School of Public Health and Health Systems, Faculty of Applied Health Sciences, University of Waterloo, Waterloo, ON, Canada

[2]Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada

[3]Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

[4]eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

Corresponding Author:
Plinio Pelegrini Morita, MSc, PEng, PhD
School of Public Health and Health Systems
Faculty of Applied Health Sciences
University of Waterloo
200 University Avenue West
Waterloo, ON, N2L 3G1
Canada
Phone: 1 5198884567 ext 31372
Fax: 1 519 746 6776
Email: plinio.morita@uwaterloo.ca

Abstract

Background: One of the main concerns of public health surveillance is to preserve the physical and mental health of older adults while supporting their independence and privacy. On the other hand, to better assist those individuals with essential health care services in the event of an emergency, their regular activities should be monitored. Internet of Things (IoT) sensors may be employed to track the sequence of activities of individuals via ambient sensors, providing real-time insights on daily activity patterns and easy access to the data through the connected ecosystem. Previous surveys to identify the regular activity patterns of older adults were deficient in the limited number of participants, short period of activity tracking, and high reliance on predefined normal activity.

Objective: The objective of this study was to overcome the aforementioned challenges by performing a pilot study to evaluate the utilization of large-scale data from smart home thermostats that collect the motion status of individuals for every 5-minute interval over a long period of time.

Methods: From a large-scale dataset, we selected a group of 30 households who met the inclusion criteria (having at least 8 sensors, being connected to the system for at least 355 days in 2018, and having up to 4 occupants). The indoor activity patterns were captured through motion sensors. We used the unsupervised, time-based, deep neural-network architecture long short-term memory-variational autoencoder to identify the regular activity pattern for each household on 2 time scales: annual and weekday. The results were validated using 2019 records. The area under the curve as well as loss in 2018 were compatible with the 2019 schedule. Daily abnormal behaviors were identified based on deviation from the regular activity model.

Results: The utilization of this approach not only enabled us to identify the regular activity pattern for each household but also provided other insights by assessing sleep behavior using the sleep time and wake-up time. We could also compare the average time individuals spent at home for the different days of the week. From our study sample, there was a significant difference in the time individuals spent indoors during the weekend versus on weekdays.

Conclusions: This approach could enhance individual health monitoring as well as public health surveillance. It provides a potentially nonobtrusive tool to assist public health officials and governments in policy development and emergency personnel

in the event of an emergency by measuring indoor behavior while preserving privacy and using existing commercially available thermostat equipment.

(JMIR Mhealth Uhealth 2020;8(11):e21209) doi:10.2196/21209

KEYWORDS

public health; IoT; anomaly detection; behavioral monitoring; deep learning; variational autoencoder; LSTM

Introduction

The Internet of Things (IoT) is a network of sensors that is integrated with physical devices and other elements to allow objects to become intelligent and interact with humans [1]. The utilization of IoT in health care is growing dramatically, especially in the areas of behavioral monitoring, welfare interventions, and incident notifications [2]. The IoT involves different types of information such as action, movements, and location, as well as physiological monitoring such as gait, heart rate, blood pressure, and stress [3]. The idea of implementing sensors to monitor one's health status and recognize activity patterns of individuals was initiated by mounting a variety of sensors on the human body [4] as well as smartphones [5] to capture different movements and other health information. In addition to wearable sensors, ambient sensors have also been utilized to build smart homes. The core feature of a smart home is activity recognition (such as watching TV, cooking meals, and sitting on the sofa) that classifies the collected data into well-defined movements [6,7]. This can be a good indicator for predicting normal and abnormal behaviors [6], as well as recognizing diseases and injuries [8,9]. As elderly people are likely to face difficulties with chronic diseases and other issues that accompany aging, a smart home could provide support and enable elderly individuals to live independently, as well as to provide immediate health care services in the event of an injury and other physical or mental health complications [2,6,10-12]. Using smart home technology would also diminish the significant burden and cost of providing long-term care services for supporting older adults that often falls on our health care systems [13-18].

In the context of smart homes, diverse approaches have been implemented to collect data from sensors and recognize different types of activities. In some studies, the activity patterns of individuals are identified by integrating wearable and ambient sensors [19,20]. In others, the interaction of humans and objects is detected through sensory-based devices [10,21,22], and the activity pattern is recognized through the sequence of those interactions [10,23].

Different supervised machine learning and deep learning models have been utilized to detect anomalies. In some approaches, participants self-labeled their activities, and normal and abnormal behaviors were identified by a supervised classification model [6]. In other studies, regular activity patterns are identified through the sequence of time-stamped events. Deviation from those patterns identified abnormal activities [22,24-27], using methods such as clustering of time-stamped events with the deep belief network feature-extraction method [21], Hidden Markov Models [22], and graph-based networks [26].

Although previous approaches towards behavioral monitoring using smart home technology were focused on collecting detailed information about health status and specific activity patterns of individuals, they ultimately proved to be obstructive, as they were limited to a small number of devices, dependent on participants to report the predefined tasks, time-consuming, and not generalizable for other smart homes [25].

Therefore, to tackle these challenges, we implemented a novel, unobstructive strategy to identify abnormal activity patterns of individuals using an ecobee smart-home thermostat. We hypothesized that, through the use of ecobee's remote sensor data, it would be possible to create models that represent typical user behaviors (eg, sleep time, wake-up time, and average time spent at home), providing the public health community with a novel tool that could identify when abnormal patterns are identified in daily user behavior. The key advantages of this approach are that we can leverage population-level data over a long period while preserving the privacy of individuals. We also implemented an unsupervised neural network for detecting anomalous activity.

Methods

Data

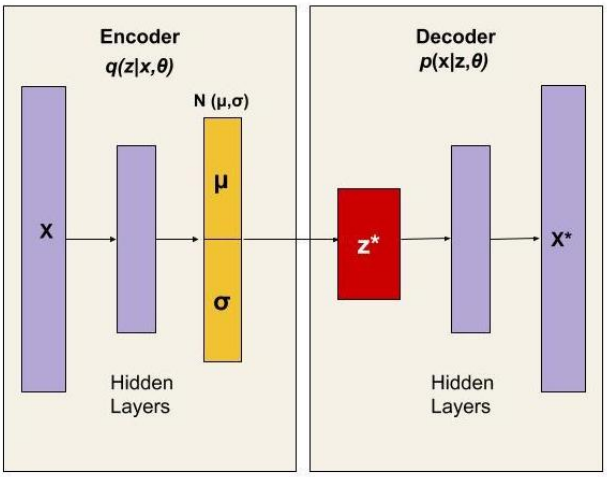
In this study, we used Donate Your Data datasets from ecobee, a smart-home thermostat manufacturer from Canada. The data are donated by households that consented to share their anonymized data to conduct research while their privacy is preserved [28]. Around 98% of the users are located in North America [29]. We selected a subset of 30 households that had at least 8 passive infrared embedded sensors and had been online for at least 355 days of 2018. The data collected by the sensors are transmitted to the thermostat base. The thermostat base can support up to 32 sensors. The data are reported in 5-minute intervals. To reduce the noise in the signal, the time window was extended from 5-minute to 30-minute intervals. A similar approach has been utilized by Huchuk et al [29] and Kleiminger et al [30]. For every 30-minute interval, motion status was determined on the basis of the sum of activation from all the sensors (if motion was captured, the sensor value changed from 0 to 1). We identified a positive motion state if the activation of at least 1 sensor lasted for 20 minutes or a minimum of 4 sensors captured a movement for a 5-minute interval each (the sum of activations during a 30-minute interval was ≥4). Therefore, each day is represented by a binary time series that identifies the motion state for the 48 time intervals during a 24-hour period.

Model

The annual record of each household was defined by a set of independent equal-sized time sequences $X = \{X1,\ldots,Xt,\ldots,XN\}$,

where each sequence is composed of 48 time units. Anomaly detection was used to decide whether the status of *Xt* was abnormal, given that all other days were known. Knowing that individuals are creatures of habit and usually have regular activity patterns, these sequences are static and periodic on a daily or weekly basis. Anomalies are the rare incidents that appear among daily patterns; therefore, identifying them through a supervised approach is challenging, as there are no labeled data available for regular and irregular activity. To answer this challenge, an unsupervised anomaly detection model should be implemented [31]. Utilizing the variational autoencoder (VAE) as a generative model can identify abnormalities by mapping the time-series data into a latent variable and reconstructing them through the latent variable [32]. In a VAE model, the encoder and decoder are defined by the probabilistic function of $q(z|x,)$ and $p(z,)$, respectively. The posterior distribution q is adapted through training and is able to map the input to a latent variable z. The variable z is assigned to a Gaussian distribution with defined parameters of mean and variance $N(\mu_x,\sigma^2_x)$. After the encoding process, the underlying characteristic of the input is generated by sampling from the Gaussian distribution (z*) that is reconstructed during the decoding process as described in Figure 1 [32].

Figure 1. Variational auto-encoder.



Due to the temporal relationship between time-series windows, the long short-term memory (LSTM)-VAE architecture was utilized to capture the latent correlation between time windows in each sequence [31]. In that way, not only are the temporal dependencies of data retained but also the information from previous time steps can be transferred to the next cell in a controlled manner. This can be done through a memory cell and 3 gates [33]. For each sequence of a time step, the hidden state (h_t) is updated by the memory cell (c_t). The memory cell is storing the information about the sequence that is controlled by the gate. The gates update the (c_t), as the proportion of current cell input (i_t), the proportion of forgetting the previous memory cell (f_t), and the proportion of current cell output (o_t) [33].

For a sample dataset X that is composed of a set of 365 independent sequences (days), each sequence was composed of 48 time intervals. The overall motion status was recorded as a unique variable in each interval. Therefore, the X dataset had the dimension of (365,48,1). Given the input X to the encoder model, the posterior distribution $q(z|x)$ was approximated by feeding the LSTM's output into 2 linear models to identify the mean and covariance of the latent variables. Then, the input of the LSTM's hidden layers from the decoder was defined by randomly sampling from the posterior distribution $q(z|x)$. The final output was defined by reconstructing the input through the random samples from the posterior distribution [34].

Based on the generative characteristics of VAE, latent variables have a key role in identifying the reconstructed data. Therefore, our objective was to model the data in such a way that the reconstructed input was similar to the original input.

This relation is represented as:

where $(X|z)$ is the distribution of generating data from the latent variable and $p(z)$ is the probability distribution of the latent variable.

On the other hand, the idea of VAE is to identify $p(z)$ using $q(z|x)$. However, identifying the distribution of $q(z|x)$ is challenging. To this end, the variational inference model is used to approximate the $q(z|x)$ distribution with simpler replacements, like standard normal or Gaussian distribution. Subsequently, the difference between the true distribution and its approximation is measured using Kullback–Leibler divergence [35].

After solving this problem, the approximation error for replacing the posterior distribution with a simpler model would be defined by:

Therefore, the overall loss function is defined as 2 parts: reconstruction error between the input and output and approximation error for replacing the posterior distribution with a simpler model.

To identify the anomaly of instances, the label and score approaches are usually used. In supervised anomaly detection, the label approach is used to identify the anomalous and normal samples using the labels 1 and 0, respectively. In contrast, for unsupervised anomaly detection, the score approach is used to identify the confidence value between 0 and 1, which reflects the likelihood of an instance being anomalous [32].

Using the VAE loss function would define the anomalous sequence on the basis of the high-loss score. In that way, the distribution of scores for all the test data points would identify the median and percentile. If the score is greater or equal than "median + IQR," it is defined as an anomalous sequence, where IQR = 75th percentile – 25th percentile.

In the training of VAE, the Keras [36] Python deep learning library was implemented. Optimization was performed using the Adam optimizer with a predefined learning rate and decay rate. The training was initialized by a mini-batch size of 7 during 200 epochs. The number of nodes for all 4 LSTM hidden layers (2 each in the encoder and decoder) was set to 100, with the *tanh* activation function for hidden layers and latent variables. We set the latent space dimensionality equal to 7 and applied an *L1* regularizer in the hidden layer of the encoder LSTM, with a weight of 0.001.

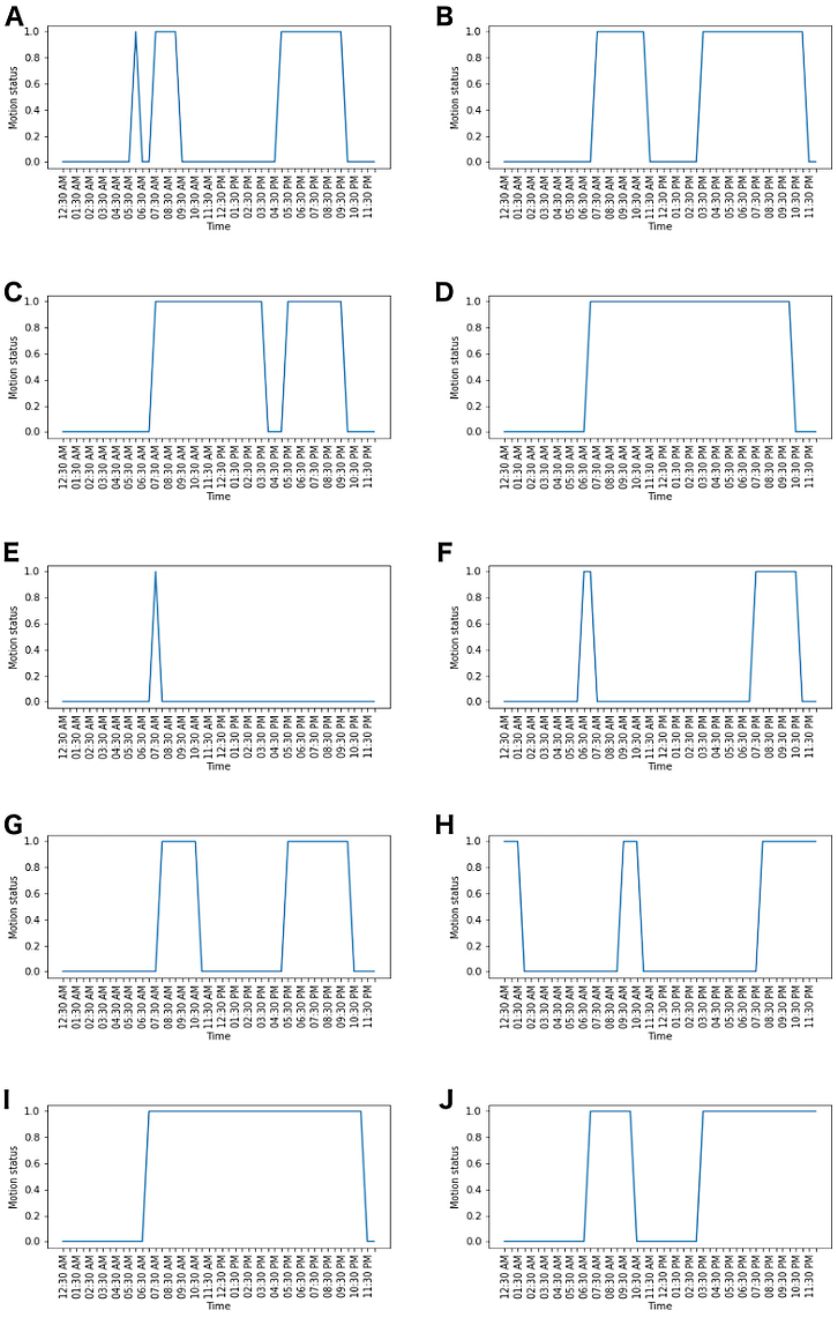
After preprocessing the data for each household, the unsupervised VAE model was implemented to identify the different regular activity patterns. From the 2018 records, the training and testing sets were randomly selected at a ratio of 80:20. The model was further validated using the records for the year 2019.

Results

Annual Pattern

The regular activity patterns for different households were identified. Figure 2 illustrates the results for the sample of 10 households. The distinct regular patterns of each household can demonstrate the diversity of the schedules.

Figure 2. Regular activity patterns for a sample of households using variational autoencoder, which can demonstrate the diverse schedules.



For household A, the wake-up time was 5:30 am. The occupants left the house at 9:30 am and returned at 4:30 pm. From 4:30 pm to 10:00 pm, indoor activity was observed, while after 10:30 pm, the absence of motion could indicate sleeping time. Similar patterns were observed for households F and G. However, for household D, the wakeup time was 6:30 am, and continuous motion was observed until 10:30 pm. Households I and C also had similar patterns. The regular activity pattern for household E demonstrated that the residents spent most of their time outside of the house.

Anomaly Detection

Anomalous activities are the rare daily patterns that differ from the regular schedule that can be defined by significant variation (median + IQR) from the regular pattern. For each household, the validation result identified the area under the curve, number of abnormal days, and average reconstruction error (loss), which are shown in Table 1.

Table 1. Validation result for each household based on the trained model.

Household ID	Abnormal days[a]	Total observed days	Loss[b]	AUC[c]	Abnormal weekend[d]
HH0	66	360	0.18	0.88	43
HH1	60	360	0.22	0.76	15
HH2	35	325	0.22	0.63	7
HH3	53	365	0.13	0.78	12
HH4	63	311	0.13	0.8	37
HH5	73	352	0.17	0.52	61
HH6	62	365	0.23	0.8	16
HH7	68	362	0.21	0.72	31
HH8	73	365	0.17	0.76	32
HH9	50	365	0.2	0.83	24
HH10	61	351	0.21	0.81	26
HH11	61	364	0.18	0.8	23
HH12	99	355	0.19	0.75	68
HH13	65	363	0.21	0.77	38
HH14	70	356	0.19	0.68	21
HH15	61	363	0.19	0.66	23
HH16	74	365	0.2	0.72	39
HH17	79	359	0.22	0.65	20
HH18	60	356	0.18	0.68	15
HH19	49	355	0.18	0.74	19
HH20	71	363	0.21	0.74	45
HH21	61	363	0.2	0.77	13
HH22	86	355	0.16	0.82	50
HH23	79	361	0.21	0.72	37
HH24	46	364	0.28	0.74	18
HH25	65	356	0.15	0.74	17
HH26	92	350	0.17	0.71	55
HH27	70	360	0.18	0.68	29
HH28	61	363	0.21	0.72	29

[a]Anomalous activity of users in 2019: activity that deviated from regular activity patterns defined from the 2018 data.

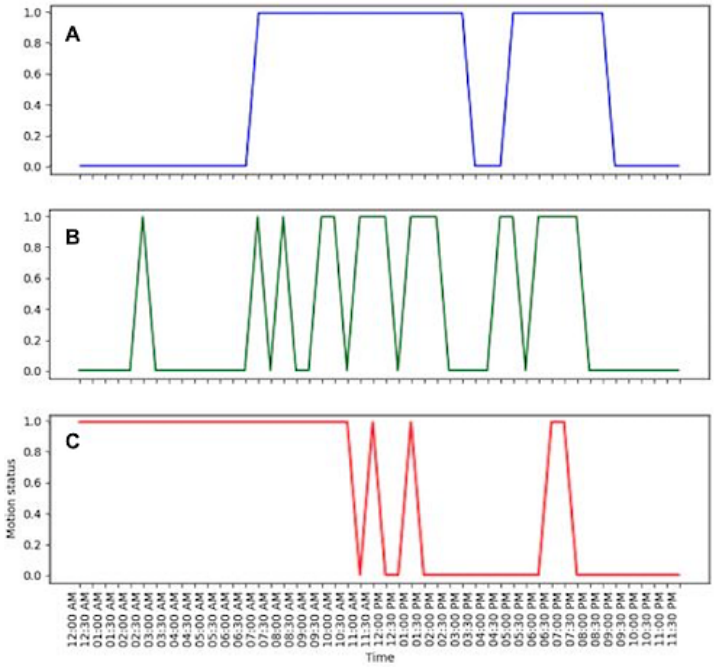
[b]Average error associated with reconstructing the validation records using the regular activity pattern.

[c]AUC: area under the curve. Overall compatibility of the regular activity pattern with validation records, in terms of recognizing the activation and deactivation of motion sensors at the right time slots.

[d]Total number of abnormal days that are weekend days.

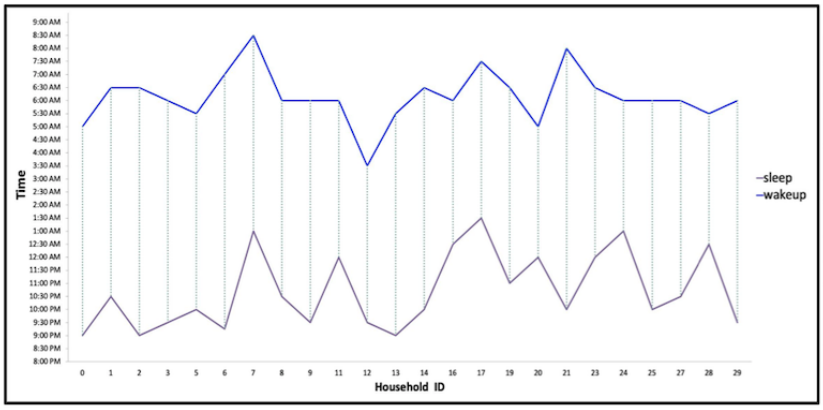
The variation in anomalous activity and normal activity with respect to the regular schedule for a sample household is illustrated in Figure 3. In the regular activity pattern of the sample household, the wake-up time was 7:00 am, and the residents spent most of the day at home. From 4:00 pm to 5:00 pm, a lack of activity is observed in the regular pattern, which could be interpreted as either the residents are usually not inside the house or they are resting or performing other sedentary activities (watching TV). A lack of activity is also observed from 9:00 pm, which can indicate sleeping time. A similar activity pattern, with a small deviation, is observed for a normal day. For a sample anomalous day, a significant deviation from the regular pattern is observed, indicating a lack of sleep during the night and a lack of activity from 2:00 pm to 6:00 pm.

Figure 3. Demonstrates the regular activity pattern for a sample household (A), reconstructed normal activity (B) and anomalous activity (C).



In addition to anomalous behavior, other indicators such as wake-up time and sleeping time may be assessed from regular activity patterns. Figure 4 demonstrates the sleep duration of households: Their wake-up time and sleep time were recognized from the regular activity pattern (Multimedia Appendix 1).

Figure 4. Sleep duration (dashed lines) for different households that had complete wake-up and sleeping time records. HH_Id: household ID.

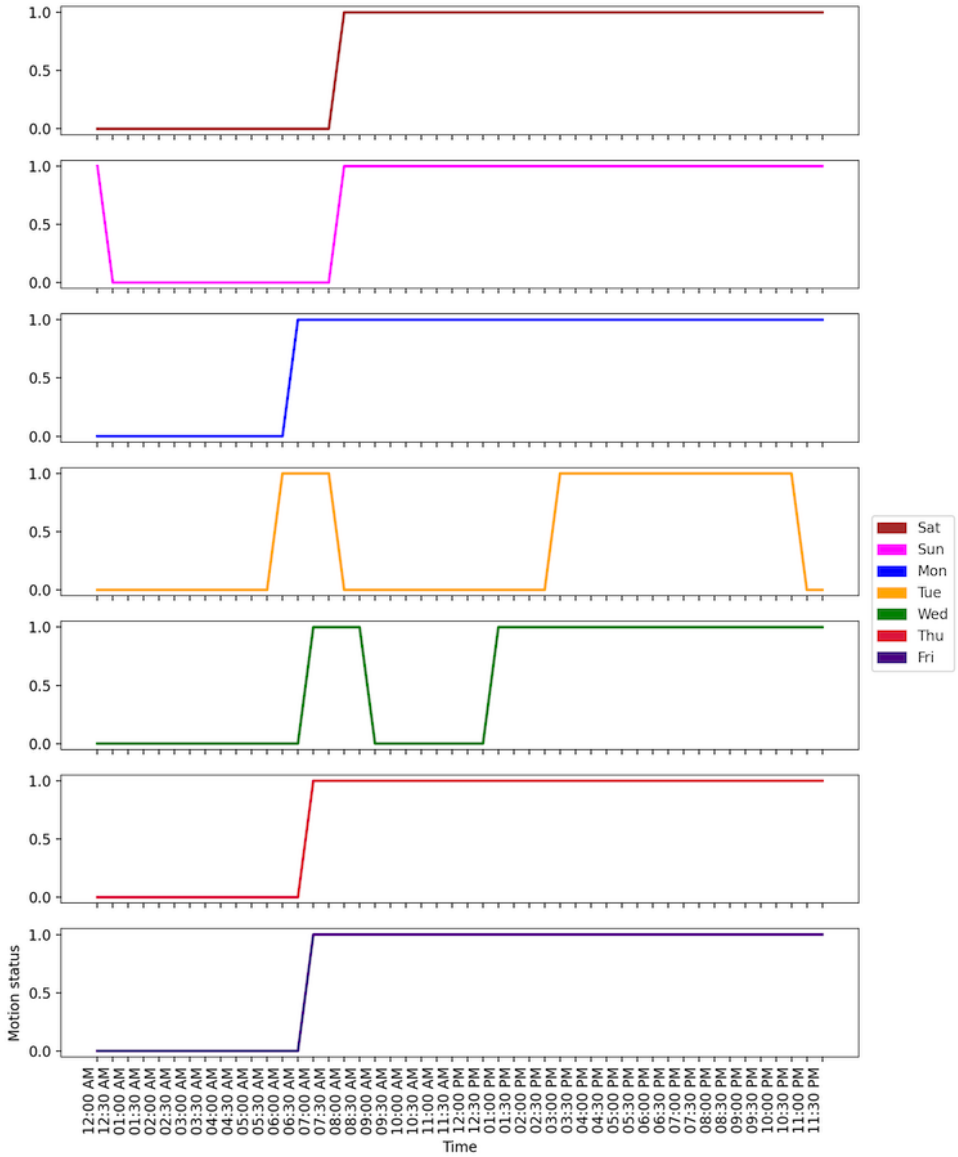


HH_Id	Sleep_duration (hours)	HH_Id	Sleep_duration (hours)
0	8	14	6
1	9.5	16	7.5
2	8.5	17	5
3	7.5	19	10
5	10	20	6.5
6	7.5	21	5
7	7.5	23	8
8	8.5	24	7.5
9	6	25	5
11	6	27	8.5
12	8.5	28	6
13	8.5	29	7.5

Weekday Patterns

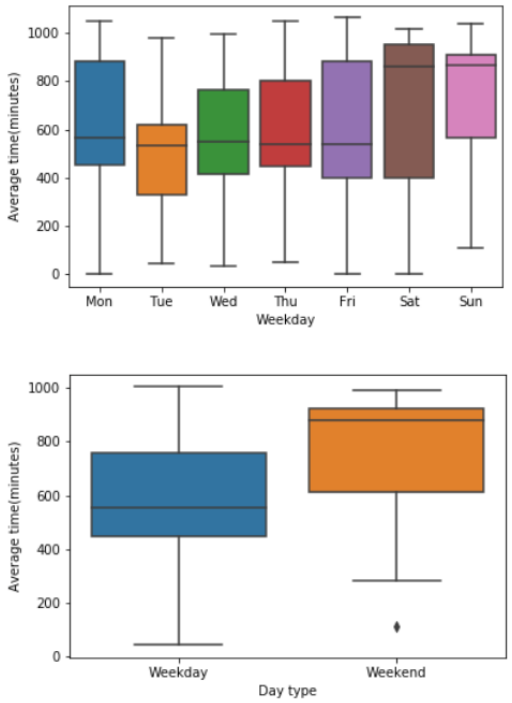
To compare the variations in regular patterns based on weekdays, the annual data were divided into subsets of weekdays, and different models were defined separately. For a sample household, the most frequent weekday pattern is defined and shown in Figure 5. Individuals had similar activity patterns from Thursday to Monday and seemed to stay home all day. However, on Tuesdays and Wednesdays, the activity pattern suggests outdoor activity during the day. The wake-up time was between 6:00 am and 7:00 am on weekdays, except Tuesdays, when the wake-up time was 5:30 am. The sleeping time was between 11:00 pm and 11:30 pm except for Sunday.

Figure 5. Weekday-specific activity pattern of a sample household.



The average number of minutes that each household spent at home was compared and is shown as box plots in Figure 6, Multimedia Appendix 2, and Multimedia Appendix 3. There was a significant increase in the time spent at home throughout the weekend versus on weekdays ($P<.05$).

Figure 6. Average minutes spent at home for (A) different days of the week and by (B) day type based on the regular activity patterns of different households.



Discussion

Principal Findings

In this study, data from the ecobee smart home thermostat was used to identify the anomalous activity patterns of individuals. The large scope of these data provides mobility recognition for every 5-minute interval. The total number of sensors in each household was different, and the locations of sensors were not identified. To enhance the accuracy of motion status, we extended the time interval from 5 minutes to 30 minutes. Since there was no specified normal activity pattern available for each household, an unsupervised LSTM-VAE method was used to generate the regular activity pattern on the basis of probabilistic distribution. From Table 1, the different AUC outcomes represent the goodness of fit of the daily record to the regular activity pattern. A low AUC value represents a lack of recognition of daily activity through the model. This could be the result of a change in the schedule (HH5). In contrast, for HH0, HH4, and HH22, the high AUC value and low loss value represent a higher chance of distinguishing the daily activity pattern through the model. The number of abnormal days was defined based on the reconstruction loss threshold (Median + IQR), which would explain the households with higher average loss and lower number of abnormal days and vice versa (eg, HH2 versus HH0).

As the demographic information of users was not specified in this study, the diverse regular activity patterns observed for each household could represent the lifestyle of working professionals. However, in the case of the older population, we are expecting a more stable schedule, and any change in behavior could be considered as a sign of unexpected incidents.

In addition to recognizing abnormal daily behavior using a regular activity pattern, other insights can be determined from the model, such as assessing sleep duration and average time spent at home. As sleep duration is one the cofounding factors of individual wellbeing, it is a major concern of public health officials and health care systems to control this risk [37]. Moreover, excessive time spent inside the house can represent the severity of detachment from the natural world and sunlight, which could have detrimental impacts such as respiratory problems or seasonal affective disorder [38]. This approach has the potential, after further validation through larger studies, to provide a nonobtrusive surveillance tool to assist public health officials and governments in policy development by reducing the public health care cost and improving the quality of services in the event of an emergency by measuring indoor behavior [13-17,39,40].

However, to enhance the interpretation of the obtained results, as well as to ascertain other outcomes such as types of incidents and identify behavioral patterns in a time-based sequence of activities, more information is required. This would include demographic information about occupants, total number of residents in each household, and location of sensors (ie, labeled by type of room in the house). In addition to monitoring the activity pattern of older adults [22,24-26], this approach could also provide insights into the impact of the COVID-19 lockdown and isolation measures on daily activity patterns of households, such as sleeping time, sleeping quality, and other indoor activities.

The outputs of the models described, when trained for each household in the dataset, has potential for supporting public health surveillance [41,42]. Using the 110,000 households available in the Donate Your Data dataset [43,44], these models could generate population-level insights on sleep patterns and indoor physical activity. Data access is one of the greatest challenges in public health research [43,44], and leveraging

available datasets such as the one used in this manuscript [14,15,18] allows an improved understanding of behavioral patterns without the added human resources necessary to collect subjective data from 110,000 households [45].

The same algorithms could be used to support independent living, by providing seniors and their families with an analytics layer that can be implemented on top of their ecobee smart thermostat technology, enabling family members to better understand the health of their loved ones; this has been previously undertaken by identifying activity recognition through postural transition using smartphones [5] as well as sensory-based devices [46] and fall detection using ambient sensors [16-18,40,47]. Ultimately, the utilization of IoT datasets such as the one provided by ecobee can provide measures of indoor activity, sleep duration, and sleep quality, as well as feedback to users in near real-time. It can also alert emergency response teams to adverse events such as elevated indoor temperatures during heatwave events [15]. The use of objective data as presented in this paper drives public health research away from subjective biases that challenge the domain [15].

Conclusion

Utilizing this approach would tackle the major challenges of public health surveillance in a more applicable and efficient way. To the best of our knowledge, this is the first study that has implemented this dataset for individual health monitoring in an unsupervised manner. The results presented in this article further the development of the UbiLab Public Health Surveillance Platform, expanding on the development of algorithms for anomaly detection.

Acknowledgments

The authors would like to thank the collaborators from ecobee for providing the Donate Your Data dataset to conduct this research. This project was funded by an NSERC Engage grant (EGP 516352-17) and an OCE VIP Grant (OCE-VIP-33080).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Average time spent home, sleeping, and awake for 30 households.
[PDF File (Adobe PDF File), 68 KB - mhealth_v8i11e21209_app1.pdf]

Multimedia Appendix 2
Average minutes spent at home for different days of the week, based on the regular activity model (minutes).
[PDF File (Adobe PDF File), 137 KB - mhealth_v8i11e21209_app2.pdf]

Multimedia Appendix 3
Comparing the average time spent home during the week versus weekend for 30 households.
[PDF File (Adobe PDF File), 73 KB - mhealth_v8i11e21209_app3.pdf]

References

1. Laplante PA, Kassab M, Laplante NL, Voas JM. Building Caring Healthcare Systems in the Internet of Things. IEEE Syst J 2018;12(3) [FREE Full text] [doi: 10.1109/JSYST.2017.2662602] [Medline: 31080541]
2. de la Torre Díez I, Alonso SG, Hamrioui S, Cruz EM, Nozaleda LM, Franco MA. IoT-Based Services and Applications for Mental Health in the Literature. J Med Syst 2018 Dec 06;43(1):11. [doi: 10.1007/s10916-018-1130-3] [Medline: 30519972]
3. Forbes G, Massie S, Craw S. Fall prediction using behavioural modelling from sensor data in smart homes. Artif Intell Rev 2019 Mar 16;53(2):1071-1091. [doi: 10.1007/s10462-019-09687-7]
4. Wan J, Gu X, Chen L, Wang J. Internet of Things for Ambient Assisted Living: Challenges and Future Opportunities. 2017 Presented at: CyberC 2017 : International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery; October 12-14, 2017; Nanjing, China p. 354-357. [doi: 10.1109/cyberc.2017.83]
5. Reyes-Ortiz J, Oneto L, Ghio A, Samá A, Anguita D, Parra X. Human Activity Recognition on Smartphones with Awareness of Basic Activities and Postural Transitions. 2014 Presented at: 24th International Conference on Artificial Neural Networks: Artificial Neural Networks and Machine Learning; September 15-19, 2014; Hamburg, Germany p. 177-184. [doi: 10.1007/978-3-319-11179-7_23]
6. Zerkouk M, Chikhaoui B. Long Short Term Memory Based Model for Abnormal Behavior Prediction in Elderly Persons. 2019 Presented at: International Conference on Smart Homes and Health Telematics: How AI Impacts Urban Living and Public Health; October 14-16, 2019; New York, NY p. 36-45. [doi: 10.1007/978-3-030-32785-9_4]

7. Biermann H, Offermann-van Heek J, Himmel S, Ziefle M. Ambient Assisted Living as Support for Aging in Place: Quantitative Users' Acceptance Study on Ultrasonic Whistles. JMIR Aging 2018 Dec 12;1(2):e11825 [FREE Full text] [doi: 10.2196/11825] [Medline: 31518245]
8. Ni Q, García Hernando AB, de la Cruz IP. The Elderly's Independent Living in Smart Homes: A Characterization of Activities and Sensing Infrastructure Survey to Facilitate Services Development. Sensors (Basel) 2015 May 14;15(5):11312-11362 [FREE Full text] [doi: 10.3390/s150511312] [Medline: 26007717]
9. Cabrita M, Tabak M, Vollenbroek-Hutten MM. Older Adults' Attitudes Toward Ambulatory Technology to Support Monitoring and Coaching of Healthy Behaviors: Qualitative Study. JMIR Aging 2019 Mar 12;2(1):e10476 [FREE Full text] [doi: 10.2196/10476] [Medline: 31518252]
10. Civitarese G, Bettini C. Monitoring objects manipulations to detect abnormal behaviors. 2017 Presented at: IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops); March 13-17, 2017; Kona, HI p. 388-393. [doi: 10.1109/percomw.2017.7917594]
11. Lotfi A, Langensiepen C, Mahmoud SM, Akhlaghinia MJ. Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour. J Ambient Intell Human Comput 2011 Jan 7;3(3):205-218. [doi: 10.1007/s12652-010-0043-x]
12. Yin J, Yang Q, Pan J. Sensor-Based Abnormal Human-Activity Detection. IEEE Trans. Knowl. Data Eng 2008 Aug;20(8):1082-1090. [doi: 10.1109/tkde.2007.1042]
13. Carnemolla P. Ageing in place and the internet of things – how smart home technologies, the built environment and caregiving intersect. Vis. in Eng 2018 Dec 6;6(1). [doi: 10.1186/s40327-018-0066-5]
14. Sahu KS, Oteomo A, Morita PP. Improving Public Health Surveillance methods via Smart Home technologies. OJPHI 2019 May 30;11(1). [doi: 10.5210/ojphi.v11i1.9795]
15. Bublitz FM, Oetomo A, Sahu KS, Kuang A, Fadrique LX, Velmovitsky PE, et al. Disruptive Technologies for Environment and Health Research: An Overview of Artificial Intelligence, Blockchain, and Internet of Things. Int J Environ Res Public Health 2019 Oct 11;16(20) [FREE Full text] [doi: 10.3390/ijerph16203847] [Medline: 31614632]
16. Alizadeh M, Shaker G, Almeida JCMD, Morita PP, Safavi-Naeini S. Remote Monitoring of Human Vital Signs Using mm-Wave FMCW Radar. IEEE Access 2019;7:54958-54968. [doi: 10.1109/access.2019.2912956]
17. Morita PP, Rocha AS, Shaker G, Lee D, Wei J, Fong B, et al. Comparative Analysis of Gait Speed Estimation Using Wideband and Narrowband Radars, Thermal Camera, and Motion Tracking Suit Technologies. J Healthc Inform Res 2020 Apr 16;4(3):215-237 [FREE Full text] [doi: 10.1007/s41666-020-00071-7]
18. Fadrique LX, Rahman D, Vaillancourt H, Boissonneault P, Donovska T, Morita PP. Overview of Policies, Guidelines, and Standards for Active Assisted Living Data Exchange: Thematic Analysis. JMIR Mhealth Uhealth 2020 Jun 22;8(6):e15923 [FREE Full text] [doi: 10.2196/15923] [Medline: 32568090]
19. Mandarić K, Skočir P, Vuković M, Ježić G. Anomaly Detection Based on Fixed and Wearable Sensors in Assisted Living Environments. 2019 Presented at: International Conference on Software, Telecommunications and Computer Networks (SoftCOM); September 19-21, 2019; Split, Croatia p. 1-6. [doi: 10.23919/softcom.2019.8903796]
20. Mehr HD, Polat H, Cetin A. Resident activity recognition in smart homes by using artificial neural networks. 2016 Presented at: 4th International Istanbul Smart Grid Congress and Fair (ICSG); April 20-21, 2016; Istanbul, Turkey p. 1-5. [doi: 10.1109/sgcf.2016.7492428]
21. Chiao HT, Basanta H, Kuo HC, Huang YP. Sensor-based detection of abnormal events for elderly people using deep belief networks. IJAHUC 2020;33(1):36. [doi: 10.1504/ijahuc.2020.10026452]
22. Ramapatruni S, Narayanan S, Mittal S, Joshi A, Joshi K. Anomaly Detection Models for Smart Home Security. 2019 Presented at: IEEE 5th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS); May 27-29, 2019; Washington, DC, USA p. 19. [doi: 10.1109/bigdatasecurity-hpsc-ids.2019.00015]
23. Chikhaoui B, Wang S, Pigot H. ADR-SPLDA: Activity discovery and recognition by combining sequential patterns and latent Dirichlet allocation. Pervasive and Mobile Computing 2012 Dec;8(6):845-862. [doi: 10.1016/j.pmcj.2012.08.004]
24. Parvin P, Paternò F, Chessa S. Anomaly Detection in the Elderly Daily Behavior. 2018 Presented at: 14th International Conference on Intelligent Environments (IE); June 25-28, 2018; Rome, Italy p. 103-106. [doi: 10.1109/ie.2018.00025]
25. Yu Y, Li C, Jonas MA, Ma C, Shezan FH, Shen S, et al. Detecting Abnormal Behaviors in Smart Home. 2019 Presented at: IEEE 16th International Conference on Mobile Ad Hoc and Sensor Systems Workshops (MASSW); November 4-7, 2019; Monterey, CA p. 37-42. [doi: 10.1109/massw.2019.00014]
26. Paudel R, Eberle W, Holder LB. Anomaly detection of elderly patient activities in smart homes using a graph-based approach. 2018 Presented at: International Conference on Data Science; Jul 30-Aug 2, 2018; Las Vegas, NV p. 163-169.
27. Aran O, Sanchez-Cortes D, Do MT, Gatica-Perez D. Anomaly Detection in Elderly Daily Behavior in Ambient Sensing Environments. 2016 Sep 22 Presented at: International Workshop on Human Behavior Understanding; October 15-19, 2016; Amsterdam, The Netherlands p. 51-67. [doi: 10.1007/978-3-319-46843-3_4]
28. Huchuk B, O'Brien W, Sanner S. A longitudinal study of thermostat behaviors based on climate, seasonal, and energy price considerations using connected thermostat data. Building and Environment 2018 Jul;139:199-210. [doi: 10.1016/j.buildenv.2018.05.003]

29. Huchuk B, Sanner S, O'Brien W. Comparison of machine learning models for occupancy prediction in residential buildings using connected thermostat data. Building and Environment 2019 Aug;160:106177. [doi: 10.1016/j.buildenv.2019.106177]
30. Kleiminger W, Mattern F, Santini S. Predicting household occupancy for smart heating control: A comparative performance analysis of state-of-the-art approaches. Energy and Buildings 2014 Dec;85:493-505. [doi: 10.1016/j.enbuild.2014.09.046]
31. Chen RQ, Shi GH, Zhao WL, Liang CH. Sequential VAE-LSTM for Anomaly Detection on Time Series Internet. 2019 Oct 10. URL: https://deeplearn.org/arxiv/96965/sequential-vae-lstm-for-anomaly-detection-on-time-series [accessed 2020-11-02]
32. Yao R, Liu C, Zhang L, Peng P. Unsupervised Anomaly Detection Using Variational Auto-Encoder based Feature Extraction. 2019 Presented at: IEEE International Conference on Prognostics and Health Management (ICPHM); June 7-20, 2019; San Francisco, CA p. 1. [doi: 10.1109/icphm.2019.8819434]
33. Pereira J, Silveira M. Unsupervised Anomaly Detection in Energy Time Series Data Using Variational Recurrent Autoencoders with Attention. 2018 Presented at: 17th IEEE International Conference on Machine Learning and Applications (ICMLA); Dec 17-20, 2018; Orlando, FL p. 1275-1282. [doi: 10.1109/icmla.2018.00207]
34. Park D, Hoshi Y, Kemp CC. A Multimodal Anomaly Detector for Robot-Assisted Feeding Using an LSTM-based Variational Autoencoder Internet. 2017. URL: http://arxiv.org/abs/1711.00614 [accessed 2020-11-02]
35. Kristiadi A. Variational Autoencoder: Intuition and Implementation. Agustinus Kristiadi's Blog. 2020. URL: https://wiseodd.github.io/techblog/2016/12/10/variational-autoencoder/ [accessed 2020-11-02]
36. Keras: the Python deep learning API. 2020. URL: https://keras.io/ [accessed 2020-11-02]
37. Hale L. Who has time to sleep? J Public Health (Oxf) 2005 Jun;27(2):205-211. [doi: 10.1093/pubmed/fdi004] [Medline: 15749721]
38. Walden S. The "Indoor Generation" and the health risks of spending more time inside. USA Today. 2018 May 15. URL: https://www.usatoday.com/story/sponsor-story/velux/2018/05/15/indoor-generation-and-health-risks-spending-more-time-inside/610289002/ [accessed 2020-11-02]
39. Sahu KS, Oetomo A, Morita P. Smart Monitoring of Population Health Risk Behaviour. 2019 Sep 15 Presented at: International Symposium on Human Factors and Ergonomics in Health Care; March 24-27, 2019; Chicago, IL p. 76-80. [doi: 10.1177/2327857919081017]
40. Boroomand A, Shaker G, Morita P, Wong A, Boger J. Autonomous gait speed estimation using 24GHz FMCW radar technology. 2018 Presented at: IEEE EMBS International Conference on Biomedical Health Informatics (BHI); March 4-7, 2018; Las Vegas, NV p. 66-69. [doi: 10.1109/bhi.2018.8333371]
41. Teutsch SM, Churchill RE, editors. Principles and Practice of Public Health Surveillance. New York, NY: Oxford University Press; 2000.
42. Henning K. Overview of Syndromic Surveillance: What is Syndromic Surveillance? MMWR Surveill Summ 2004 Sep 24;53(Suppl):5-11.
43. Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. PLoS Comput Biol 2015 Feb;11(2):e1003904 [FREE Full text] [doi: 10.1371/journal.pcbi.1003904] [Medline: 25664461]
44. Ola O, Sedig K. The challenge of big data in public health: an opportunity for visual analytics. Online J Public Health Inform 2014;5(3):223 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4933] [Medline: 24678376]
45. Mooney SJ, Pejaver V. Big Data in Public Health: Terminology, Machine Learning, and Privacy. Annu Rev Public Health 2018 Apr 01;39:95-112 [FREE Full text] [doi: 10.1146/annurev-publhealth-040617-014208] [Medline: 29261408]
46. Suryadevara NK, Mukhopadhyay SC. Wireless Sensor Network Based Home Monitoring System for Wellness Determination of Elderly. IEEE Sensors J 2012 Jun;12(6):1965-1972. [doi: 10.1109/jsen.2011.2182341]
47. Yazar A, Erden F, Cetin AE. Multi-sensor ambient assisted living system for fall detection. 2014. URL: https://www.semanticscholar.org/paper/45faf7d1b208af47177ba21a34f916c658232690 [accessed 2020-11-02]

Abbreviations

AUC: area under the curve
IoT: internet of things
LSTM: long short-term memory
VAE: variational autoencoder

Edited by G Eysenbach, L Buis; submitted 08.06.20; peer-reviewed by D Istrat, M Görges; comments to author 29.06.20; revised version received 27.07.20; accepted 24.10.20; published 13.11.20.

Please cite as:
Jalali N, Sahu KS, Oetomo A, Morita PP
Understanding User Behavior Through the Use of Unsupervised Anomaly Detection: Proof of Concept Using Internet of Things Smart Home Thermostat Data for Improving Public Health Surveillance
JMIR Mhealth Uhealth 2020;8(11):e21209
URL: http://mhealth.jmir.org/2020/11/e21209/
doi:10.2196/21209
PMID:33185562

Original Paper

Enabling Remote Patient Monitoring Through the Use of Smart Thermostat Data in Canada: Exploratory Study

Kirti Sundar Sahu[1], MPH, MPT; Arlene Oetomo[1], BSc; Plinio Pelegrini Morita[1,2,3,4,5], PhD, PEng

[1]School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada

[2]Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada

[3]Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

[4]eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

[5]Research Institute for Aging, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Plinio Pelegrini Morita, PhD, PEng
School of Public Health and Health Systems
University of Waterloo
200 University Avenue West
Waterloo, ON, N2L 3G1
Canada
Phone: 1 5198884567 ext 31372
Fax: 1 519 746 6776
Email: plinio.morita@uwaterloo.ca

Abstract

Background: Advances in technology have made the development of remote patient monitoring possible in recent years. However, there is still room for innovation in the types of technologies that are developed, used, and implemented. The smart thermostat solutions provided in this study can expand beyond typically defined features and be used for improved holistic health monitoring purposes.

Objective: The aim of this study is to validate the hypothesis that remote motion sensors could be used to quantify and track an individual's movements around the house. On the basis of our results, the next step would be to determine if using remote motion sensors could be a novel data collection method compared with the national census-level surveys administered by governmental bodies. The results will be used to inform a more extensive implementation study of similar smart home technologies to gather data for machine learning algorithms and to build upon pattern recognition and comprehensive health monitoring.

Methods: We conducted a pilot study with a sample size of 8 to validate the use of remote motion sensors to quantify movement in the house. A large database containing data from smart home thermostats was analyzed to compare the following indicators; sleep, physical activity, and sedentary behavior. These indicators were developed by the Public Health Agency of Canada and are collected through traditional survey methods.

Results: The results showed a significant Spearman rank correlation coefficient of 0.8 ($P<.001$), which indicates a positive linear association between the total number of sensors activated and the total number of indoor steps traveled by study participants. In addition, the indicators of sleep, physical activity, and sedentary behavior were all found to be highly comparable with those attained by the Public Health Agency of Canada.

Conclusions: The findings demonstrate that remote motion sensors data from a smart thermostat solution are a viable option when compared with traditional survey data collection methods for health data collection and are also a form of zero-effort technology that can be used to monitor the activity levels and nature of activity of occupants within the home.

(JMIR Mhealth Uhealth 2020;8(11):e21016) doi:10.2196/21016

KEYWORDS

disruptive technology; health information systems; public health surveillance; health behavior; Internet of Things; cell phone; mobile phone

Introduction

Background

Remote patient monitoring (RPM) involves digital technologies used to collect and transmit medical and health data from individuals to health care providers for assessments and recommendations [1,2]. As a component of public health surveillance, novel RPM technology infrastructures can modernize traditional, time-consuming, and expensive data collection methods [1]. The Internet of Things (IoT) enables health care providers to remotely monitor patient health and analyze the data with minimal delay to develop a personalized treatment plan or follow the patient's progress over time [3]. Aspects that can be monitored include physical activity levels, drug adherence, and physiological indicators such as blood pressure and heart rate [1,4]. The ultimate goal of RPM is to enable individuals to lead healthier lives and improve their well-being at home by leveraging technology.

Although a variety of RPM platforms exist today [5], consumers continue to be limited by interoperability issues because the systems lack compatibility, often creating silos of technologies where data cannot be integrated or exchanged. This study uses existing off-the-shelf technologies, such as IoT-based smart thermostats equipped with wireless motion sensors and consumer-level wearable fitness trackers. Smart Wi-Fi thermostats regulate indoor air temperature via motion sensors that detect occupancy. Their benefits include cost and energy savings owing to increased heating and cooling efficiency and their ability to eliminate the need for an additional dedicated home monitoring platform. In the United States, customers using these smart thermostats save approximately 23% on heating and cooling costs [6]. For this study, the Ubiquitous Health Technology Lab (UbiLab) [7] at the University of Waterloo has partnered with ecobee, a Toronto-based Canadian smart Wi-Fi thermostat company [8]. Using data from ecobee thermostats and a variety of sensors, the UbiLab Public Health Surveillance Platform (UPHSP) was developed to augment current public health surveillance efforts through the use of IoT data. Public health surveillance is labor intensive, requires significant human resources, often relies on outdated data, and has numerous types of biases (eg, recall, performance, nonresponse, voluntary response, social desirability) [9]. The UPHSP improves that scenario by providing access to near real time, objective sensor-based data.

Public health surveillance is the systematic collection, analysis, and interpretation of health data that are required to determine funding for programs, strategies, and initiatives [10]. It is also an important component of public health that focuses on early identification, mitigation of disease, prevention, planning, and overall evaluation of population health behaviors. At present, the Canadian public health surveillance system relies on self-reporting and short-term activity monitoring to collect data on health indicators included in the Physical Activity, Sedentary Behavior and Sleep (PASS) Indicator framework [11]. This method has implications for research methodology, the validity of research results, and the soundness of public policy developed from evidence using questionnaire-based research [12]. Traditional methods have several limitations. First, they are prone to recall and social desirability bias. Second, data collection does not occur in real time and requires extensive administrative resources adding to the total costs [13]. Third, conducting surveys can be a lengthy process and thus are done periodically and intermittently. Finally, methods such as physical activity monitoring require participants to be completely involved in data collection, which may be inconvenient and cause deviations from their regular routines, leading to data collection challenges that result in smaller study sizes [14].

Objectives

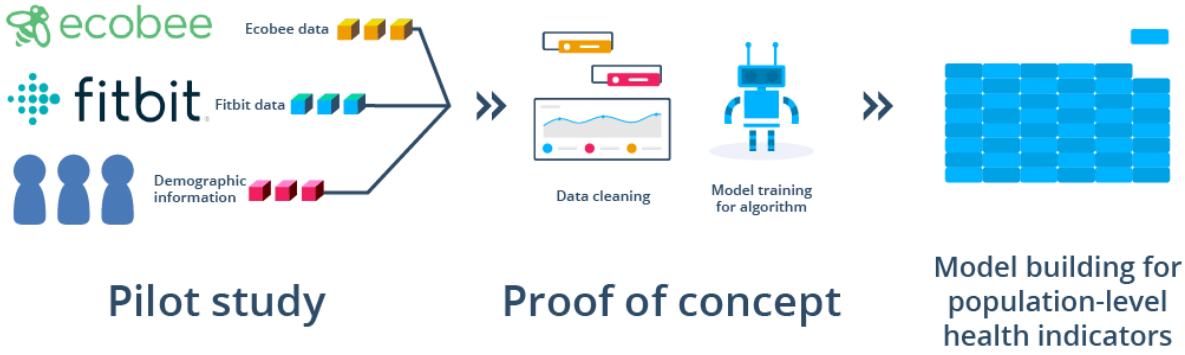
The goal of this study is to develop and facilitate the use of UPHSP across Canada for population- and individual-level surveillance of health behaviors. The platform enables RPM via smart Wi-Fi thermostats and motion sensors to seamlessly collect and transmit health behavior data. These data are provided to public health officials, care providers (with minimum disturbance of daily activities and delivery of personalized health insights), and users. UPHSP has advantages over conventional data collection methods because it addresses current challenges with traditional data collection techniques. UPHSP allows researchers to access granular and longitudinal data generated directly from within the participant's house. The anonymized data will be collected 24 hours a day and 7 days a week and then consolidated so that they are available to health care officials in near real time. The core strength of this surveillance method is that it utilizes zero-effort technology (ZET) [14] that requires no effort from users. This effectively reduces the burden on study participants, ensures minimal disturbance to their daily routines, and can also be applied in the real-world environment [15].

In collaboration with the Public Health Agency of Canada, UbiLab is developing UPHSP to improve and develop individual- and population-level health behavior indicators. In the initial stages of the project, the team focused on the existing set of health indicators from the PASS Indicators framework. PASS Indicators measured from within the house were selected for this study. This study was then submitted to the Healthy Behaviour Data Challenge organized by the Public Health Agency of Canada (PHAC), the Canadian Institutes for Health Research, and the MaRS Discovery District in August 2017 [16,17]. The aim was to demonstrate that collecting reliable data for the 3 indicators was possible with smart home technology. The goal is to enable public health officials, such as the PHAC, to access real time population-level data of Canadian behaviors and develop policies and programs with relevant data to improve the health and well-being of Canadians.

Methods

UPHSP is based on thermostat sensor technology that collects raw motion data generated by activity in the house. A two-step approach was carried out to achieve the objectives of this study. The first step involved deploying a small pilot study to validate the accuracy of the motion sensors, and the second step investigated the larger data set acquired via the ecobee's Donate Your Data (DYD) program [18,19] to demonstrate the scalability of this study, as described in Figure 1.

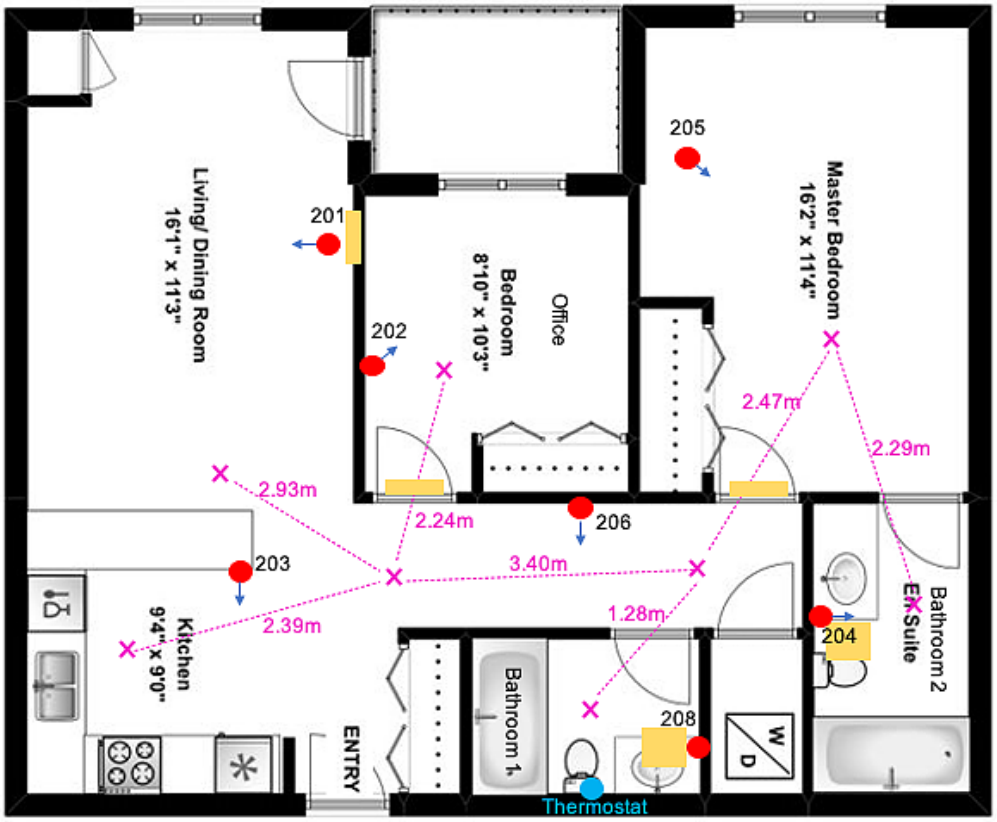
Figure 1. The process of model building for population-level health surveillance.



Pilot Study

The first important task for this study was to determine if the ecobee sensors were a valid data source. The association between remote motion sensors and a person's movement in the house needed to be definite and accurate. A pilot study was conducted to validate the relationship between ecobee remote sensor activation and movement within the house and to test the accuracy of these sensors. The team hypothesized that there would be a significant association between the number of sensors activated and the number of steps taken by the participants in the house. The study included a total of 8 participants (4 women and 4 men) aged between 25 and 41 years and were residents of the Kitchener-Waterloo Region in Ontario, Canada. Each participant wore a Fitbit Zip and was the sole occupant in their house during the data collection period. Furthermore, each house was equipped with an ecobee thermostat unit, and remote sensors were placed around the house to ensure maximum coverage by the sensors as seen in the Figure 2. Approximately 5 to 30 remote motion sensors were deployed in each home, depending on the size of the house. The floorplan layout for each house was obtained to determine optimum sensor placement. Motion and step data were collected in all 5 houses via the thermostat sensors and the Fitbit between 9:00 AM and 5:00 PM. Participants were also instructed to record their activities at all times and identify interruptions, such as using the washroom or getting a snack, in an activity log.

Figure 2. Sample home layout of a participant with ecobee thermostat and remote sensors.



Once the data were collected, cleaned, and ready for analysis, the team was able to identify the activity status of the participants (active vs sedentary) using step data and remote sensor activations (Figure 1). As the data were not normally distributed, the Spearman correlation test was the tool of choice to test for association [20]. The Spearman correlation coefficient was found to be 0.8, with $P<.001$ indicating that there was a strong positive linear association between the total number of

sensors activated and the total number of steps traveled by the participant within the house and that this correlation was statistically significant. The strength of this correlation increased with the increasing number of sensors within a larger house. With our first hypothesis successfully proven, the second phase of the project, as identified in Figure 1, could begin.

DYD

Ecobee created the *Donate Your Data* (DYD) [21,22] program to provide researchers access to anonymized data collected from their technology for further energy and sustainability research. UbiLab, since its launch in 2016, is the first research team to use these data for health purposes and has gained access to a large data repository through this research partnership. The first data set received from ecobee contained over 10,000 records (or households) and has since increased to more than 110,000 records after 3 years.

The metadata information accompanying the thermostat motion data is provided by the ecobee user upon initial setup of their ecobee thermostat, where they may choose to opt into the program to contribute their data to the DYD program. The information that is collected includes the style of house (eg, apartment, townhouse, etc), floor square footage, number of floors, and number of occupants. Incomplete fields were eliminated from the analysis pool. For ecobee to maintain this program's opt-in process and encourage participation, leaving the fields optional increased the enrollment but did not contribute to completeness or quality of the data set. The DYD data set is completely anonymized to maintain user privacy, and all user customizations of sensor names or locations are not shared with researchers. This is a limitation but a sensible move from ecobee to protect user privacy and encourage program participation. Unfortunately, the metadata quality is restricted by the amount and quality of the information entered during setup by the user.

From January 2015 to March 2017, data from 556 single occupant houses in the United States and 70 in Canada were included in the analysis. According to ecobee, over 1 million smart thermostats are installed in houses across the United States and Canada and that number continues to grow. Furthermore, federal and provincial rebate programs encourage user adoption of smart thermostat technologies [23]. These efforts are beneficial for this study and have future potential in UPHSP.

A Comparison

The team focused on the ability to collect data from sensors and demonstrate how the same indicators could be measured with sensor technology. Existing health indicators from the PASS Indicator framework were matched on the basis of the measure description determined by the PHAC [11]. At present, the sleep indicator is a self-reported value determined by the Canadian Health Measures Survey (CHMS) from Statistics Canada and is the sum of the average number of hours spent sleeping in a 24-hour period by adults aged 18 to 79 years [24]. This is potentially problematic because the value is subject to recall bias and sleep times are rounded to the nearest half hour. In addition, the platform provides insight into the sleep timing, patterns, and number of disturbances captured owing to excessive movement or additional triggering of sensors in areas of the house other than the bedroom. Indicators of sleep interruptions would provide insights into sleep-related conditions and disorders such as sleep apnea and sleepwalking, but they have not yet been developed by the PHAC. However, the PHAC is currently developing several new sleep indicators for sleep quality, sleep hygiene, and more.

Adherence to the PASS Indicators of physical activity guidelines and the total amount of moderate-to-vigorous physical activity is derived from CHMS and uses accelerometer data [25]. As the platform functions within the house, the lack of movement during the day very likely indicates that the individual was out of the house. Canada has distinct seasons; therefore, this information could be used to investigate seasonal patterns that contribute to social isolation or increased healthy behavior during the warmer months.

Physical Activity

Ecobee data are streamed in 5-min intervals, which means that every 5 min the remote sensors send the motion data to the thermostat and the thermostat sends these data to the server. Hence, all the analyses performed in the project were performed in 5-min intervals. The total activity within the house was measured for the purposes of our study. Physical activity was defined as true if ≥2 sensors were activated at 5-min intervals within the house during the waking period between 8:00 AM and 10:00 PM. This method of measuring physical activity provides real time data, removes the involvement of the user from data collection, effectively eliminates recall bias, and minimizes the social desirability effect.

Sedentary Behavior

Sedentary behavior was defined as less than 2 sensors being activated within any 5-min interval during the day (between 8:00 AM and 10:00 PM) while the individual was in the house. As this was a pilot study, we tried to classify unknown data into 3 categories: sleep, sedentary behavior, and physical activity. Our assumptions were as follows: (1) zero sensor activation was labeled as sleep; (2) 3 or more activated sensors indicated physical activity; and (3) the remaining values indicated sedentary behavior. The logic behind this type of sampling was that if an individual moved from one place to another within 5 min, at least two sensors would be activated and the energy equivalent used would be less than when undertaking physical activity. When more than 2 sensors were activated, the step values were more appropriately deemed to be a higher level of physical activity.

This is different from the PHAC accelerometer-generated indicator obtained from CHMS [26] because the focus was on sedentary time within the house as our technology did not measure sedentary time spent commuting or while at work. The algorithms developed for the project treated movement between rooms as physical activity and only stagnant behavior was considered sedentary.

Results

Pilot Study

Initially, a pilot study was conducted to confirm the accuracy of the responses of the ecobee sensors to movement in the house. It involved using Fitbit to track the steps of 8 participants for approximately one week, which was equivalent to 386 person-hours of data.

A total of 8 participants used Fitbit for 1 week; however, after processing and cleaning the data, not all of the data were exactly of same length across the 7 days. There was no significant difference between the duration of participants' data used. A simple visual inspection of the number of steps captured by the Fitbits and sensor activations was done to determine alignment (as seen in Figures 3 and 4). There is a clear increase in motion sensor activation in alignment with increased steps taken. To confirm that the sensors were working precisely, a statistical test was performed to measure the association between the sensor activation and Fitbit data. The Spearman correlation coefficient was r=0.8 (range 0.78-0.90; n=3292; $P<.001$). These results indicate a strong association between sensor activation and the number of steps recorded via Fitbit. Figures 3 and 4 illustrate the sensors that were activated when an individual walked around their house and their association with steps taken by that individual captured via Fitbit. The scatterplot in Figure 5 demonstrates the data points and individual-level correlation between steps and activated ecobee sensors. In addition, traveling between rooms, which is indicated by more steps, showed higher levels of activity on the sensors. During periods of inactivity, when no steps were recorded by Fitbit, the sensors also displayed no or minimal activation. A few random or unexplained activations may be caused by internal movements of items in the house, such as a moving curtain or a bird flying by a window that faced a sensor. These movements were considered noise and removed for the purposes of our analysis. Only one sensor was activated during this time, which can be attributed to disturbed sleep or other kinds of sedentary behavior.

Figure 3. Association between Fitbit and ecobee database–separated.

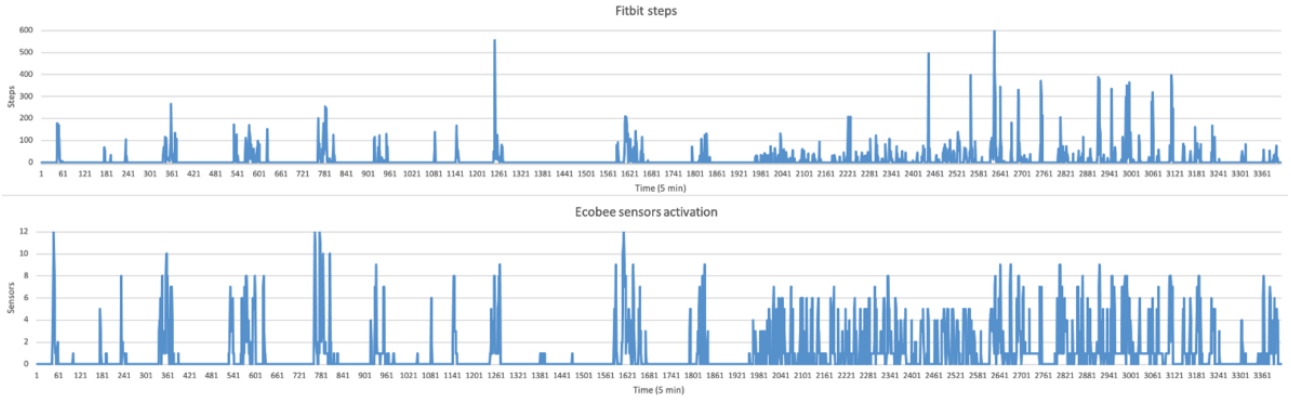


Figure 4. Association between Fitbit and ecobee database–superimposed.

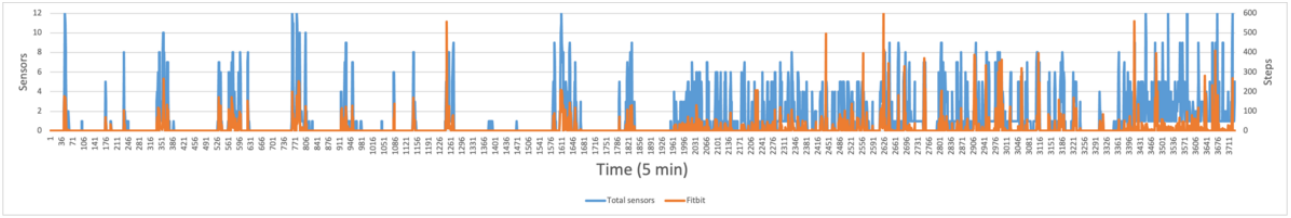
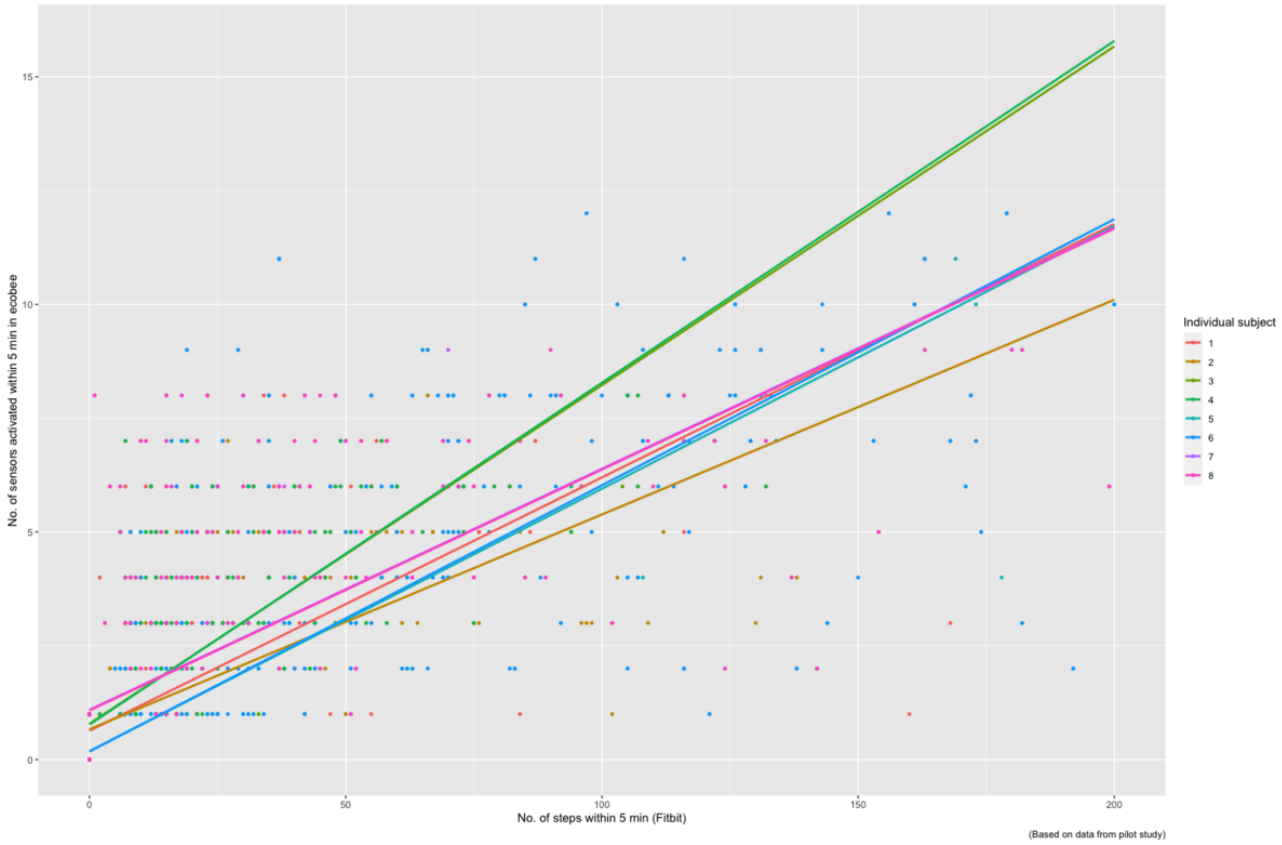


Figure 5. Scatter plot with individual correlation.



Sleep

Following the pilot study, the second phase of the project involved analyzing the larger DYD data set from ecobee. Sleep in households was measured using UPHSP. In Table 1, the individual and household columns are defined as follows: a house with a single occupant was labeled as *individual* and households with more than one occupant was labeled as *household*. They were mutually exclusive. The findings were that, on average, single occupant households from the ecobee data set had a sleep duration of 7.8 hours per day. Compared with the conclusions from previous surveys and PHAC's current method of measuring average Canadian sleep duration [24], the platform's result was 0.7 hours higher on average. UPHSP can calculate the number of sleep interruptions during the night using automated algorithms. Data recorded by ecobee revealed approximately two hours of interrupted sleep during regular sleep by ecobee single occupancy residents. An interruption was defined as any movement within the sleeping hours. Interrupted sleep is a new indicator as it is not currently possible to self-report this measure and has not been collected or reported by the PHAC.

Table 1. Key findings and a comparison of UbiLab findings with PHAC reported physical activity, sedentary behavior, and sleep indicators (n=958).

Name of indicators	UbiLab[a]		PASS[b]
	Individual (n=70)	Household (n=888)	Individual
Nighttime sleep, hours	7.89	7.71	7.2
Disturbed sleep, hours	2.10	2.28	N/A[c]
Physical activity in the house, minutes per day	85.2	146.4	24.1
Sedentary time, hours	4.44	5.75	9.6
Away period, hours	8.12	5.80	N/A

[a]UbiLab: Ubiquitous Health Technology Lab.

[b]PASS: physical activity, sedentary behavior, and sleep.

[c]N/A: not applicable.

Physical Activity

The second indicator measured by UPHSP was the amount of physical activity in each household, which averaged 85 min per day. Physical activity was defined as movement within the house during waking hours. A comparison of our indicator with physical activity (moderate to vigorous) measurements reported by the PHAC revealed that UPHSP results surpassed that of

PHAC by almost an hour (85 min vs 24 min). However, our measurements do not consider the intensity of the physical activity (ie, moderate or vigorous). Further studies are required to develop intelligent algorithms that can differentiate between the intensities of physical activity.

Analyzing the data over a long period can inform public health officials about seasonal physical activity variations. During the winter months, the number of indoor activities and the time spent inside will likely increase owing to cold weather and shorter days. Measuring household physical activity in Canada is therefore essential and necessary from a public health perspective.

Sedentary Behavior

The study participants recorded their activity log during the pilot study to track the type and duration of their activities. The platform measures general activities in the house, time patterns of doing certain types of housework, the number of interruptions, and the area of the house where the users spend most of their time. The log depicts the total time spent on a certain activity that is recorded by a participant. In the first row, the time the user spent cleaning their house would be considered physical activity as the user was active and moved around the house. In this case, multiple sensors would have been active throughout the house during a 5-min interval. The second row indicates that the user was sitting and watching television. This would be considered a sedentary behavior if less than 2 sensors were activated in a 5-min interval. The number of interruptions is a new indicator that could help the PHAC understand healthy behaviors and seasonal patterns.

UPHSP calculates the true nature and timing of all the behaviors being carried out in the house. In addition, the data set was filtered for any exceptional behavior to ensure that there was an accurate analysis of the ecobee users' average activities. For this purpose, 100 steps (156 m) were selected as the threshold for a 5-min interval. If participants took more than 100 steps in a 5-min interval (eg, while working out on a treadmill), these occurrences were considered outliers. To select this set point of 100 steps, the Spearman correlation statistical test was performed on several thresholds beginning with 100 steps, which was increased by 50 steps at a time to a maximum of 400 steps. As seen in the Multimedia Appendix 1, there was only a small amount of variation in the correlation coefficients, $r \geq 0.798$ and $P < .001$, at 100 steps, which indicates a strong correlation between the distance traveled and the activation of sensors. Therefore, the project was successfully able to use ecobee remote sensors to understand indoor physical activity levels.

Discussion

Principal Findings

Technological advancements have rapidly become cross-disciplinary as they merge health care and medicine. However, unlike the retail and commercial sectors, the health care field lags behind [25]. Mobile health is the application of digital technology for the use of medical care and has the goal of empowering individuals to be in control of their own health. Technology has revolutionized the field of medicine and continues to impact all areas of health care. Medical technology has evolved rapidly, enabling doctors to use new equipment inside hospitals and allowing them to connect with patients and other physicians thousands of kilometers away through smart devices [27]. Hundreds of health and wellness mobile apps have been developed since the introduction of smartphones (also known as the field of mobile health [mHealth]) [28]. These apps enable users to track their own health (ie, fitness levels, water consumption, diet, etc) and compare their data with standardized guidelines or challenge their peers to step competitions. The dependence on medical technology cannot be ignored within the health care industry because health care professionals can continue to find ways to improve their practice with the development of these brilliant innovations [27].

Manipulating the temperature of a house using a smartphone was not possible 10 years ago. At present, monitoring health and diagnosing conditions are possible via smart home technologies. With this technology, UbiLab is turning valuable sources of data into useful insights. Raw data are difficult to interpret and would be insufficient to encourage behavioral change. They are made interpretable with artificial intelligence (AI) algorithms using data from the pilot study. Interpreted health data will be shared with users using mobile or web-based platforms and will be presented through a dashboard. These data could help individuals learn about their daily activity patterns, general health behaviors, and compare their behaviors with those of their peers or the national average. For this to become a reality, it is necessary to ensure that all the users can view and interpret the data anywhere with clarity. A simple dashboard interface will be carefully designed to display meaningful and actionable health information for Canadians. The team created a simple dashboard for the purpose of the Healthy Behaviour Data Challenge. Ideally, such a dashboard would display the user's health data in a textual, auditory, and visual manner. It would support multiple languages and also be interactive and visually appealing to gain more attention from individuals. Furthermore, the dashboard would recommend changes to the user's physical activity, sleep, and sedentary behaviors to improve their overall health. Smart home technologies are currently expanding into the realm of virtual assistants that are based on AI such as Amazon Echo [29], Apple's Siri [30], and Google Home [31]. Incorporation of these technologies could bring additional personalization to the platform and user interaction and further motivate health behavior change and enhance user experience.

UPHSP has shown potential by providing a possible improvement and by increasing the efficiency of the current labor-intensive methods of data collection. In comparison with traditional data collection methods that require labor, funding, and ample time, this platform is more cost-effective because it leverages existing technology (smart thermostats). The platform is very convenient for most users because it uses ZET and only requires direct involvement for the initial setup and orientation of the system. In addition, the data collected via UPHSP is more granular and accurate, collected 24/7 in real time, which eliminates potential social desirability and recall bias. On average, the current delay in research and policy development is 17 years [27,32,33], which can be changed with the help of

technology and multisite studies. Implementing UPHSP could significantly reduce the time spent on data entry and collection because the technology requires little active involvement from study participants and can reduce the overall study time.

The implementation of this study is possible because of a growing interest in smart home technology. An estimated 100,000 Canadian households have an ecobee thermostat installed, with even more users in the United States of America, and this number is only expected to increase. Incentives from the provincial government, such as the Ontario Green Fund, has also helped drive interest and adoption. Similar initiatives would diversify the sample population, which is valuable as it would open up this technology to a wider sociodemographic of population. The issue of a biased sample population can be addressed by investing in the adoption of smart thermostat across the country. By including individuals from different cultures, ages, and various socioeconomic statuses, UPHSP can help the PHAC gain insight into contrasting health behaviors among different segments of the country. This could help identify populations with the greatest need and inform policies that target these groups.

There are many potential applications of this type of technology and ways of leveraging of nontraditional data sources, as has been done here with data from IoT for public health surveillance [34]. For example, public health officials can monitor risky behavior for chronic diseases, and this could have a significant impact on routine monitoring systems at the national and provincial levels. Real time monitoring of population health behaviors without interfering with daily activity is the biggest strength of this study. As technology rapidly evolves to become more affordable, smaller, and wireless, passive and seamless data collection for health monitoring becomes more feasible.

Limitations

The team has identified some limitations with the technology that constrain the capabilities of the data collected and the possible health behavior insights. The ecobee remote sensors collect data at 5-min intervals, which from the perspective of collecting data for temperature control is sufficient. However, a higher level of granularity is required for RPM, and the system is not yet able to differentiate between multiple, different occupants in a house. There is no means to tag an individual to identify them as they move through the house; therefore, observing a pattern is difficult, and it is not yet possible to attach it to a specific person. Identifying or differentiating occupants in the house is currently challenging without adding an additional on-body device. The larger DYD data set is a biased sample because the typical buyer of the smart thermostat is a middle-to-upper class individual who is interested in cost and energy savings, reducing environmental impact, and technologically savvy. Initiatives from governments and smart technology manufacturers to encourage the adoption of smart thermostat technologies, such as rebates or the installation of this technology in affordable housing units, can help to make this more accessible to a wider income range and reduce the sample bias. The distribution will also need to adequately reflect the Canadian population, and considerations should be taken to include smaller or remote communities outside of the provinces of Ontario and British Columbia.

Moving forward, the UbiLab will collect demographic information from a subset of existing ecobee users to understand the association between age, sex, and other relevant demographic indicators. Our plan is to explore other sensor technologies to train machine learning algorithms and generate data. In addition, solutions to correctly identify true positive presence in the house will be explored to address this shortcoming.

Conclusions

There is a lot of potential for RPM to expand and leverage commercial technologies. This study is just one example where technology can be used to bring innovative solutions for real use in the realm of health care, especially as it allows the use of technologies that are zero effort and have more than one added benefit. Technologies such as these will be able to advance the fields of RPM and public health surveillance.

Acknowledgments

UbiLab would like to thank all study participants for their time and the volunteers of UbiLab and ecobee for their support in the completion of this study. The authors would like to thank Dr. Shannon Majowicz and Dr. Joel Dubin from the University of Waterloo for their guidance for this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Spearman correlation coefficients for each of the data cleaning stages.
[DOCX File , 14 KB - mhealth_v8i11e21016_app1.docx]

References

1. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote Patient Monitoring via Non-Invasive Digital Technologies: A Systematic Review. Telemed J E Health 2017 Jan;23(1):3-17 [FREE Full text] [doi: 10.1089/tmj.2016.0051] [Medline: 27116181]

2. Rojahn K, Laplante S, Sloand J, Main C, Ibrahim A, Wild J, et al. Remote Monitoring of Chronic Diseases: A Landscape Assessment of Policies in Four European Countries. PLoS One 2016;11(5):e0155738 [FREE Full text] [doi: 10.1371/journal.pone.0155738] [Medline: 27195764]
3. Dimitrov DV. Medical internet of things and big data in healthcare. Healthc Inform Res 2016 Jul;22(3):156-163 [FREE Full text] [doi: 10.4258/hir.2016.22.3.156] [Medline: 27525156]
4. 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]
5. Clarke M, de Folter J, Verma V, Gokalp H. Interoperable End-to-End Remote Patient Monitoring Platform based on IEEE 11073 PHD and ZigBee Health Care Profile. IEEE Trans. Biomed. Eng 2020:1-1. [doi: 10.1109/tbme.2017.2732501]
6. Savings from your ecobee. Ecobee. URL: https://www.ecobee.com/savings/
7. Towards Zero-effort Remote Monitoring. University of Waterloo. 2018. URL: https://uwaterloo.ca/ubiquitous-health-technology-lab/ [accessed 2020-09-30]
8. Ecobee. URL: https://www.ecobee.com/ [accessed 2018-08-28]
9. Rosenman R, Tennekoon V, Hill LG. Measuring bias in self-reported data. Int J Behav Healthc Res 2011;2(4):320. [doi: 10.1504/ijbhr.2011.043414] [Medline: 25383095]
10. Universal Health Coverage. World Health Organization. URL: http://www.who.int/topics/public_health_surveillance/en/ [accessed 2018-08-28]
11. Physical Activity, Sedentary Behaviour, and Sleep Indicator Framework. Public Health Infobase - Canada. URL: https://infobase.phac-aspc.gc.ca/pass-apcss/ [accessed 2018-08-28]
12. Bowling A. Mode of questionnaire administration can have serious effects on data quality. J Public Health (Oxf) 2005 Sep;27(3):281-291. [doi: 10.1093/pubmed/fdi031] [Medline: 15870099]
13. Schootman M, Nelson E, Werner K, Shacham E, Elliott M, Ratnapradipa K, et al. Emerging technologies to measure neighborhood conditions in public health: implications for interventions and next steps. Int J Health Geogr 2016 Jun 23;15(1):20 [FREE Full text] [doi: 10.1186/s12942-016-0050-z] [Medline: 27339260]
14. Mihailidis A, Boger J, Hoey J, Jiancaro T. Zero effort technologies: considerations, challenges, and use in health, wellness, and rehabilitation. Synt Lect Assist Rehab Health-Preserv Technol 2011 Aug 30;1(2):1-94. [doi: 10.2200/s00380ed1v01y201108arh002]
15. van Hoof J, Kort H, Rutten P, Duijnstee M. Ageing-in-place with the use of ambient intelligence technology: perspectives of older users. Int J Med Inform 2011 May;80(5):310-331 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.02.010] [Medline: 21439898]
16. Healthy Behaviour Data Challenge Winners Announced. Canada Health Infoway: Digital Health in Canada. 2018. URL: https://www.infoway-inforoute.ca/en/what-we-do/blog/innovation/7818-healthy-behaviour-data-challenge-winners-announced [accessed 2018-08-28]
17. The Healthy Behaviour Data Challenge. HeroX: Crowdsourcing Platform for Enterprise and Innovators. URL: https://www.herox.com/healthdatachallenge/community [accessed 2018-08-13]
18. Social Impact. Ecobee. URL: https://www.ecobee.com/social-impact/ [accessed 2018-08-13]
19. CNET. URL: https://www.cnet.com/news/why-ecobee-created-donate-your-data-program-despite-privacy-worries-q-a/ [accessed 2018-08-28]
20. Tully MA, McBride C, Heron L, Hunter RF. The validation of Fibit Zip physical activity monitor as a measure of free-living physical activity. BMC Res Notes 2014 Dec 23;7:952 [FREE Full text] [doi: 10.1186/1756-0500-7-952] [Medline: 25539733]
21. Donate Your Data. Ecobee. 2016. URL: https://www.ecobee.com/donateyourdata/ [accessed 2020-11-11]
22. Ecobee Room Sensor Data to Help Advance Healthcare. Ecobee. URL: https://www.ecobee.com/2017/02/ecobee-room-sensor-data-to-help-advance-healthcare/ [accessed 2018-11-11]
23. Low Carbon Economy Fund: Energy Savings Rebate program. Natural Resources Canada. URL: https://www.canada.ca/en/environment-climate-change/services/climate-change/low-carbon-economy-fund/energy-savings-rebate.html [accessed 2020-11-11]
24. Physical Activity, Sedentary Behaviour and Sleep (PASS) Indicators. Public Health Infobase - Canada. URL: https://health-infobase.canada.ca/pass/indicator-details/1147 [accessed 2020-11-11]
25. Canadian Health Measures Survey (CHMS) Data User Guide: Cycle 1. Statistics Canada. URL: http://www23.statcan.gc.ca/imdb-bmdi/pub/document/5071_D2_T1_V1-eng.htm#a6_4_7 [accessed 2018-11-12]
26. Measure Description. Public Health Infobase - Canada. URL: https://infobase.phac-aspc.gc.ca/pass-apcss/indicator-details-en.aspx?id=31 [accessed 2018-08-28]
27. Brownson R, Kreuter M, Arrington B, True W. Translating scientific discoveries into public health action: how can schools of public health move us forward? Public Health Rep 2006;121(1):97-103 [FREE Full text] [doi: 10.1177/003335490612100118] [Medline: 16416704]
28. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3 [FREE Full text] [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]

29. Echo (2nd Generation) - Smart Speaker With Alexa and Dolby Processing - Charcoal Fabric. Amazon. URL: https://www.amazon.com/all-new-amazon-echo-speaker-with-wifi-alexa-dark-charcoal/dp/B06XCM9LJ4 [accessed 2018-08-28]
30. Siri. Apple. URL: https://www.apple.com/ca/ios/siri/ [accessed 2018-08-28]
31. Google. 2018. URL: https://store.google.com/product/google_home [accessed 2018-08-28]
32. Morris Z, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. J R Soc Med 2011 Dec;104(12):510-520. [doi: 10.1258/jrsm.2011.110180] [Medline: 22179294]
33. Munro C, Savel R. Narrowing the 17-year research to practice gap. Am J Crit Care 2016 May;25(3):194-196. [doi: 10.4037/ajcc2016449] [Medline: 27134218]
34. Bublitz FM, Oetomo A, Sahu KS, Kuang A, Fadrique LX, Velmovitsky PE, et al. Disruptive technologies for environment and health research: an overview of artificial intelligence, blockchain, and internet of things. Int J Environ Res Public Health 2019 Oct 11;16(20):- [FREE Full text] [doi: 10.3390/ijerph16203847] [Medline: 31614632]

Abbreviations

AI: artificial intelligence
CHMS: Canadian Health Measures Survey
DYD: Donate Your Data
IoT: Internet of Things
PASS: Physical Activity, Sedentary Behavior and Sleep
PHAC: Public Health Agency of Canada
RPM: remote patient monitoring
UbiLab: Ubiquitous Health Technology Lab
UPHSP: UbiLab Public Health Surveillance Platform
ZET: zero-effort technology

Edited by G Eysenbach; submitted 08.06.20; peer-reviewed by D Istrat, R Lundin; comments to author 24.07.20; revised version received 30.07.20; accepted 22.09.20; published 20.11.20.

Please cite as:
Sahu KS, Oetomo A, Morita PP
Enabling Remote Patient Monitoring Through the Use of Smart Thermostat Data in Canada: Exploratory Study
JMIR Mhealth Uhealth 2020;8(11):e21016
URL: https://mhealth.jmir.org/2020/11/e21016
doi:10.2196/21016
PMID:33216001

Original Paper

Using a Mobile App-Based International Classification of Functioning, Disability, and Health Set to Assess the Functioning of Spinal Cord Injury Patients: Rasch Analysis

Mengmeng Jia[1*], BSN; Jie Tang[2*], BSN; Sumei Xie[3], BSN; Xiaokuo He[4], PhD; Yingmin Wang[5], MSN; Ting Liu[1], MSN; Tiebin Yan[5], PhD; Kun Li[1], PhD

[1]School of Nursing, Sun Yat-sen University, Guangzhou, China

[2]Department of Spinal Cord Injury Rehabilitation, Sichuan Provincial Rehabilitation Hospital, Chengdu, China

[3]Department of Spinal Cord Injury Rehabilitation, Guangdong Provincial Work Injury Rehabilitation Hospital, Guangzhou, China

[4]Department of Rehabilitation Medicine, The Fifth Hospital of Xiamen, Xiamen, China

[5]Department of Rehabilitation Medicine, Sun Yat-sen Memorial Hospital, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Kun Li, PhD
School of Nursing
Sun Yat-sen University
No. 74 Zhong Shan Second Road
Guangzhou,
China
Phone: 86 138 22206519
Fax: 86 020 87333043
Email: likun22@mail.sysu.edu.cn

Abstract

Background: The International Classification of Functioning, Disability, and Health (ICF) is a unified system of functioning terminology that has been used to develop electronic health records and assessment instruments. Its application has been limited, however, by its complex terminology, numerous categories, uncertain operationalization, and the training required to use it well. Together is a mobile health app designed to extend medical support to the families of spinal cord injury (SCI) patients in China. The app's core framework is a set of only 31 ICF categories. The app also provides rating guidelines and automatically transforms routine assessment results to the terms of the ICF qualifiers.

Objective: The goal of the research is to examine the suitability of the ICF set used in the app Together for use as an instrument for assessing the functioning of SCI patients.

Methods: A cross-sectional study was conducted including 112 SCI patients recruited before discharge from four rehabilitation centers in China between May 2018 and October 2019. Nurses used the app to assess patient functioning in face-to-face interviews. The resulting data were then subjected to Rasch analysis.

Results: After deleting two categories (family relationships and socializing) and one personal factor (knowledge about spinal cord injury) that did not fit the Rasch model, the body functions and body structures, activities and participation, and contextual factors components of the ICF exhibited adequate fit to the Rasch model. All three demonstrated acceptable person separation indices. The 28 categories retained in the set were free of differential item functioning by gender, age, education level, or etiology.

Conclusions: Together overcomes some of the obstacles to practical application of the ICF. The app is a reliable assessment tool for assessing functioning after spinal cord injury.

(JMIR Mhealth Uhealth 2020;8(11):e20723) doi:10.2196/20723

KEYWORDS

International Classification of Functioning, Disability and Health; spinal cord injuries; mobile health app; Rasch analysis

Introduction

The International Classification of Functioning, Disability, and Health (ICF) is a unified system of terminology for multidisciplinary use issued in 2001 by the World Health Organization (WHO). It provides a consensus framework for defining functioning and disability and their interrelationships with health conditions and contextual factors [1]. As a standard language, the ICF is designed to be easily understood and used among multidisciplinary teams [2]. The comprehensive perspective on functioning and interdisciplinary focus have motivated the development of data collection tools, electronic health records, and assessment instruments [3-5], and it is now sometimes viewed as a third health indicator for monitoring a health system's performance after mortality and morbidity [6]. Many studies have provided evidence of the ICF's value in reflecting patient levels of functioning, helping decision making, enhancing collaboration, and planning treatment.

However, challenges have limited practical application of the ICF, including its relatively complex terminology and category numbering. Each ICF category has its own distinct definition, which doesn't always accord with the prevailing medical terminology. Professionals need to be trained before using the system [5]. Although many ICF core sets with fewer categories have been specifically developed for certain conditions, some studies still report that the application of the ICF is time-consuming [7]. Reducing the number of categories by selecting only the most relevant remains challenging. Additionally, rating using the ICF qualifiers is not easy. There are 5 grades: 0 = no problem, 1 = mild, 2 = moderate, 3 = severe, and 4 = complete problem. But there can be large differences among user assessments because of a lack of clear assessment guidelines, giving the approach poor interrater reliability [8-11]. Some researchers have suggested establishing assessment guidelines, but these would necessarily be complex and require additional user training, further hindering the system's acceptance [12,13].

A mobile health app is a health-related software program installed on smart mobile devices that provides health information and tracks a user's health behavior. It might also allow remote consultation [14]. With the popularity of personal mobile devices such as cellphones and tablets, mobile apps have been used in many fields in health care [15,16]. Studies have shown that app-based transitional care enhances patients' self-perceptions of efficacy [17], helps prevent complications [18,19], improves quality of life [18], and reduces readmissions to hospital [18].

The study team previously developed an ICF-based app called Together for the transitional care of spinal cord injury (SCI) patients in China [20]. SCI is a serious and life-changing disease that can cause paraplegia or quadriplegia. Most SCI patients in China live at home after their acute treatment and rehabilitation [21]. They almost always need professional medical support for further rehabilitation and preventing complications [22,23]. However, community medical resources in China are at present limited and cannot meet SCI patients' complex long-term health needs [24]. Together was designed to bring professional health care support from medical institutions to SCI patients living with their families in China. The language of the app is Chinese, and the copyright is held by China's Sun Yat-sen University [25].

Together's core framework is a set of ICF categories that reflect the levels of functioning typical of SCI patients and help organize online assessment, standardize health guidance, and coordinate interdisciplinary collaboration. The app uses fewer and more specific categories than the normal ICF core set, focusing on the transitional care of SCI patients. Preliminary studies identified 31 ICF categories as the most useful outcome indicators in the transitional care of SCI patients, covering the major physiological, psychological, and social participation problems of SCI patients at home [20,26]. The app provides consistent assessment prompts for the raters. Guidelines are provided for rating each ICF category, which helps to guarantee the consistency of the assessments, minimize the training required, and ease the load on the clinical staff doing the assessments.

This study was part of a research program designed to document the effects of an app-based transitional care model for SCI patients at home. Rasch analysis was applied to examine the suitability of the app's set of categories. The overall aim was to determine to what extent Together can solve problems related to using the ICF in clinical practice.

The assumption of the Rasch model is that a person with greater ability is more likely than a person with less ability to pass in relation to an item, and that an easy item is more likely to be passed than a difficult one [27]. A Rasch analysis is used to examine whether an instrument makes those distinctions satisfactorily. A person's performance on an item should be related only to the person's ability and the difficulty of the item, regardless of gender, age, education, etc. [27]. If an instrument fits the Rasch model well, it can be used to reflect the performance of people with different abilities. This study was designed to test Together's performance in that regard. A good result could present a new approach in the use of the ICF.

Methods

Study Design

A cross-sectional design was employed involving four research centers in Guangzhou, Chengdu, and Shiyan in China. The study was approved by Sun Yat-sen University's ethics committee (file 2017ZSLYEC-0620).

Participants

The participants were recruited between May 2018 and October 2019 prior to discharge from the four research centers. The inclusion criteria were age 18 years or older, SCI according to the International Standards for Neurological Classification of Spinal Cord Injury [28] and imaging examination, less than 2 years since injury, conscious and able to communicate, and in possession of an internet-connected mobile device and familiarity with using it. The exclusion criteria were severe heart, brain, lung, liver, or kidney disease; acute-stage spinal cord injury or in the critical period; or spinal cord lesions with a degenerative, genetic, or congenital cause.

App

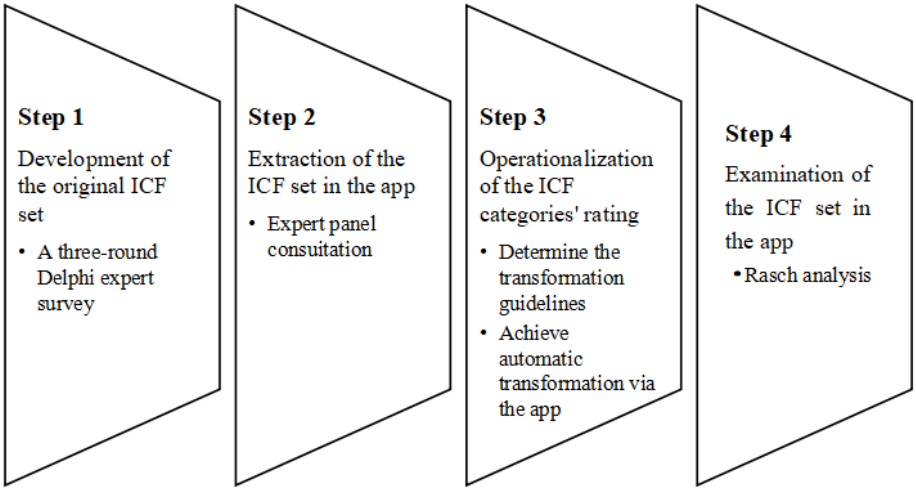
The Together app was designed to help medical staff provide remote follow-up for home-dwelling SCI patients. Hospital-based nurses, physicians, and therapists responsible for the transitional care for SCI patients at home are the target users. The core functions of the app comprise online assessment, providing standard health guidance, interdisciplinary referral within the team, interaction among health staff and patients, and management of online follow-up. With the help of the app, health care personnel can assess patient performance in terms of ICF categories and remotely provide health education to patients according to the assessment results. The app can be used to refer patients to different professionals on the health care team. Weekly reminders make managing follow-up by medical staff easier.

The app's development has been reported previously [20,26]. From 51 ICF categories identified as outcome indicators useful in the transitional care of SCI patients by experts in the field via a 3-round Delphi survey, 31 were selected by a panel of 5 experts to form the app's framework based on the feasibility of use in clinical practice. them. These categories best reflected the dysfunctions and complications most common and most in need of monitoring for SCI patients at home. Categories reflecting actual performance of SCI patients in daily life were preferred.

The 31 ICF categories address physiological functioning, psychological functioning, complications, daily living activities, social participation, adaptation to environmental factors, and personal factors. For each category, guidelines were established for converting routine clinical assessment results to the ICF qualifiers, and a standardized guidance program was formulated by the expert panel based on the knowledge-attitude-practice theory. Together is an Android app developed using the Java language. Most of the app's functions—online assessment, providing standard health guidance, interdisciplinary referral—apply to all of the categories.

The app's utility depends heavily on to what extent its assessment results reflect functioning differences among patients with different capabilities. This study used Rasch analysis to test the app's suitability as an assessment instrument. Figure 1 shows the development and examination process.

Figure 1. Development process of the Together app.



Assessment Guidelines

Medical staff communicate with the patients face to face or by telephone, assess them on each ICF category, and record the results in the app, which offers standard verbal prompts to the clinician to unify different clinicians' assessments with respect to each ICF category. For example, the verbal prompt of the sensation of pain (b280) category was "If 0 is not painful and 10 is the most painful, how serious is your pain?" Four transformation guidelines were developed to transform the initial clinical assessment results: 0 = no problem, 1 = mild, 2 = moderate, 3 = severe, or 4 = complete problem (Table 1) [29]. The clinician is responsible for the initial assessments and the app automatically transforms their initial assessment results to the ICF qualifiers according to the preset guidelines.

Guideline 1 transforms patient information in the form of percentages to the ICF qualifiers. Using muscle power functions (b730) as an example, the 0 qualifier would indicate that all of the key muscles below the injured neurological level had power grade >3, and the 4 qualifier would indicate that 95% to 100% of key muscles had power grade <3. Guideline 2 transforms the wording of patient reports to the qualifiers. Using mobility of joint functions (b710) as an example, no limitation of joint mobility, slight limitation, moderate limitation, severe limitation, and total immobility would be rated as 0, 1, 2, 3, and 4, respectively. Guideline 3 transforms the frequency with which a problem was observed to the qualifiers. Using increased blood pressure (b4200) as an example, stable blood pressure over the past month would be rated as 0, whereas high blood pressure almost every day would receive a 4. Guideline 4 transforms assessment results generated using routine clinical instruments or standards, such as the 0-10 numerical rating scale (NRS) for pain, to the qualifiers. NRS scores of 0, 1-2, 3-4, 5-9, and 10 would be rated as 0, 1, 2, 3, and 4, respectively (Multimedia Appendix 1).

Table 1. Guidelines for transforming routine assessment results to the International Classification of Functioning, Disability, and Health qualifiers.

ICF[a] qualifier	Guideline 1[b]	Guideline 2[c]	Guideline 3[d]	Guideline 4[e]
0 (no problem)	0%-4%	No, none, absent, negligible...	The person has no such problem.	NRS[f] for pain, MAS[g], NPIAP[h] stage, FIM[i], WHOQoL-BREF[j], SF-36[k]
1 (mild problem)	5%-24%	Mild, slight, low...	The problem rarely happened in the last month (<25% of the time).	NRS[f] for pain, MAS[g], NPIAP[h] stage, FIM[i], WHOQoL-BREF[j], SF-36[k]
2 (moderate problem)	25%-49%	Moderate, medium, fair...	The problem happened occasionally in the last month (<50% of the time).	NRS[f] for pain, MAS[g], NPIAP[h] stage, FIM[i], WHOQoL-BREF[j], SF-36[k]
3 (severe problem)	50%-95%	Severe, high, extreme...	The problem happened frequently in the last month (>50% of the time).	NRS[f] for pain, MAS[g], NPIAP[h] stage, FIM[i], WHOQoL-BREF[j], SF-36[k]
4 (complete problem)	96%-100%	Complete, total...	The problem happened almost every day in the last month (>95% of the time).	NRS[f] for pain, MAS[g], NPIAP[h] stage, FIM[i], WHOQoL-BREF[j], SF-36[k]

[a]ICF: International Classification of Functioning, Disability, and Health.

[b]Guideline 1: transforms patient information in the form of percentages to the ICF qualifiers.

[c]Guideline 2: transforms wording from patient reports to the ICF qualifiers.

[d]Guideline 3: transforms the frequency with which a problem was observed during the previous month to the ICF qualifiers.

[e]Guideline 4: transforms the scores of a routine clinical instrument or standards to the ICF qualifiers.

[f]NRS: numeric rating scale.

[g]MAS: Modified Ashworth Scale.

[h]NPIAP: National Pressure Injury Advisory Panel.

[i]FIM: Functional Independence Measure.

[j]WHOQoL-BREF: World Health Organization Quality of Life Assessment–Abbreviated.

[k]SF-36: 36-Item Short Form Health Survey.

Instruments

Demographic and Disease Questionnaire

The questionnaire consisted of two parts, including demographic and disease-related data such as name, gender, age, education level, diagnosis, etiology, American Spinal Injury Association Impairment Scale grade, SCI level, and duration of the disability. The information was collected by nurses in face-to-face interviews with the patients and by reviewing their medical records.

App-Based International Classification of Functioning, Disability, and Health Set

The 31 categories came from the three components body functions and body structures (15 categories), activities and participation (10 categories), and contextual factors (6 categories). Although the ICF does not classify personal factors in the contextual factors component, four personal factor items related to the psychology of SCI patients were included based on preliminary testing (acceptance of life in a wheelchair/in bed, knowledge about spinal cord injury, coping with everyday life, and adjustment to new body image) [20,26]. The app was then used to rate the SCI patients' functioning on a 1-4 scale. In addition, an option 9 (not applicable) was also used. For the activities and participation component, each patient's performance on tasks in actual life situations (not in a standard environment) was assessed. For the contextual factors component, barriers (but not facilitators) encountered in each patient's life were assessed. In previous research, the eating (d550) and drinking (d560) categories were found to be strongly interrelated and difficult to assess separately [13]; therefore, those categories were combined into a testlet to be assessed.

Data Collection

All nurses who participated in the study first received half a day of training involving a lecture and workshop, including an introduction to the study and how to use the app. Eligible patients were invited to participate in the study before discharge. After signing the informed consent form, demographic and disease-related data were collected by the trained nurses in face-to-face interviews and by reviewing the patients' medical records. The nurses then assessed the patients' performance with respect to each ICF category using the app. They did this face to face referring to the standard verbal prompts for each category in the app. The app system allows submission only after all categories have been evaluated; otherwise, the system indicates to the user that the evaluation is incomplete.

Analysis

Data Processing

SPSS Statistics software version 21.0 (IBM Corporation) was used to analyze the demographic and disease data. RUMM2030 software (RUMM Laboratory Pty) was used to perform the Rasch analysis. For each component of the ICF set, the overall fit to a Rasch model was examined. If the overall fit was not good, poorly fitting categories were identified and deleted.

Another round of Rasch analysis was then run until adequate overall fit was attained. The following properties of the ICF set were examined.

Overall Fit to the Rasch Model

A nonsignificant value in a χ^2 test for item-trait interaction, a mean within ±2.5 (SD <1.5) for the fit residuals of the items and persons indicate good overall fit to the Rasch model [30]. Fit residuals represent the extent to which the observations do not fit a Rasch model. The significance level was adjusted using the Bonferroni correction [31].

Single-Item Fit to the Rasch Model

The good fit of a single category was represented by a nonsignificant χ^2 test and a mean of the fit residual values within ±2.5 [32].

Person Separation Index

An acceptable person separation index indicates good internal consistency for the instrument and reflects the ability of the instrument to discriminate between people with different abilities. It has a range of 0 to 1, with higher values indicating a better ability (>0.7 indicates good) [33].

Differential Item Functioning

For an ideal Rasch model, no factors should influence a person's performance regarding an item except the Rasch factors [30]. The differential item functioning shows the influences of the other factors. In this study, a nonsignificant analysis of variance result was taken as indicating no differential item functioning for a specific category based on gender, age, education level, or etiology.

Results

Patient Characteristics

The demographic and disease characteristics of all 112 spinal cord injury patients are shown in Table 2. Their ages ranged from 18 to 65 years (mean 41.7 [SD 12.3]); 82.1% (92/112) of the patients were younger than 60 years, with patients in the 40- to 49-year age group the most numerous; 83.0% (93/112) of the patients were male. A total of 60.8% (68/112) claimed to have had a middle school education and 25.9% (29/112) only primary education or less. The duration of their disability ranged from 1 to 22 months (mean 7.1 [SD 4.2]); 88.4% (99/112) had been injured for less than a year. Most of the injuries (100/112, 89.3%) were caused by trauma. Most of the patients were injured at the thoracic and cervical levels, accounting for 50.0% (56/112) and 27.7% (31/112), respectively. About half (57/112, 50.9%) of the patients had complete injury.

Table 2. Characteristics of the study sample (n=112).

Characteristic	Value, n (%)
Gender	
Male	93 (83.0)
Female	19 (17.0)
Age in years	
18-29	27 (24.1)
30-39	19 (17.0)
40-49	36 (32.1)
50-59	20 (17.9)
60-65	20 (17.9)
Education	
Primary school and below	29 (25.9)
Junior high school	47 (42.0)
Senior high school	21 (18.8)
College and above	15 (13.4)
Etiology	
Trauma	100 (89.3)
Nontrauma	12 (10.7)
Duration of disease in months	
1-6	59 (52.7)
7-12	40 (35.7)
13-18	12 (10.7)
19-22	1 (0.9)
American Spinal Injury Association Impairment scale	
Complete injury	57 (50.9)
Incomplete injury	55 (49.1)
Spinal cord injury level	
Cervical	31 (27.7)
Thoracic	56 (50.0)
Lumbar sacral	25 (22.3)

Rasch Analysis Results

The 31 ICF categories belonged to body functions and body structures (15), activities and participation (10), and contextual factors (6). To attain adequate fit to the Rasch model for each component, categories that did not fit were deleted as multiple rounds of Rasch analysis were conducted. Table 3 shows the process and results of the Rasch analysis for each component.

Table 3. Summary of results of the Rasch analyses (n=112).

Analysis and action	Item fit residual, mean (SD)	Person fit residual, mean (SD)	Overall model fit[a]		Person separation index
			χ^2	*P* value	
Body functions and body structures					
1 Original categories	–0.41 (0.86)	–0.32 (0.62)	41.5	.08	0.50
Activities and participation					
2 Original categories	–0.48 (2.27)	–0.38 (0.80)	75.7	<.001[b]	0.89
3 Deleted family relationships (d760)	–0.18 (2.37)	–0.30 (0.79)	43.3	<.001[b]	0.89
6 Deleted family relationships (d760) and socializing (d9205)	–0.06 (1.59)	–0.23 (0.73)	24.7	.08	0.89
Environmental factors and personal factors					
7 Original categories	0.25 (1.63)	–0.45 (1.30)	32.8	.001[b]	0.65
8 Deleted knowledge about spinal cord injury	–0.33 (1.32)	–0.43 (1.22)	13.6	.19	0.68

[a]Overall model fit was tested using a χ^2 test with a Bonferroni-adjusted *P* value. The values were all *P*<.01.

[b]Significant according to the Bonferroni-adjusted *P* value.

In the first-round Rasch analysis, the body functions and body structures component consisting of 15 categories exhibited a nonsignificant χ^2 test result for the item-trait interaction (χ^2_{30}=41.5, *P*=.08, Bonferroni-adjusted *P*=.05/15=.003). Additionally, the means for the item and person fit residuals were within ±2.5 (SD <1.5). These results suggested a good fit to the Rasch model. All of the 15 categories exhibited nonsignificant χ^2 test results and the means of their fit residual values were also within acceptable limits (Table 4). The person separation index of this component was 0.5. There was no differential item functioning for any of the categories by gender, age, education level, or etiology.

Regarding the initial activities and participation component with 10 categories, in the first-round Rasch analysis, the χ^2 test for the item-trait interaction yielded a significant result (χ^2_{20}=75.7, *P*<.001, Bonferroni-adjusted *P*=.05/10=.005). The single-item fit analysis found that four categories, changing basic body position (d410), washing oneself (d510), family relationships (d760), and socializing (d9205), did not fit the Rasch model. The family relationships (d760) and socializing (d9205) categories both exhibited poor fit results in a previous study [13]. Considering the mean fit residual values 3.253 for family relationships (d760) and 2.605 for socializing (d9205) and the category meanings, d760 was deleted first. However, the χ^2 test result for the item-trait interaction in the second-round Rasch analysis (after deleting d760) remained significant (χ^2_{18}=43.3, *P*<.001, Bonferroni-adjusted *P*=.05/9=.0056). A third round of Rasch analysis was performed after deleting both family relationships (d760) and socializing (d9205). Although the standard deviation of the overall item fit residuals (SD 1.59) was a little larger than the upper limit, the χ^2 test result for the item-trait interaction was no longer significant (χ^2_{16}=24.7, *P*=.08, Bonferroni-adjusted *P*=.05/8=.0063), suggesting good fit to the Rasch model. The single-item fit tests for the remaining 8 categories also yielded good model fit results, with nonsignificant χ^2 test results (Table 4). The person separation index of the component was excellent (0.89) and no differential item functioning was detected for any of the categories by gender, age, education level, or etiology.

For the contextual factors component, the first-round Rasch analysis starting with 6 categories indicated poor model fit according to the χ^2 test result for the item-trait interaction (χ^2_{12}=32.8, *P*<.001, Bonferroni-adjusted *P*=.05/6=.0083). The following single-item fit analysis showed that the personal factor knowledge about spinal cord injury did not fit well (χ^2_2=19.1, *P*<.001). After deleting that item, the component displayed satisfactory overall model fit, with a nonsignificant χ^2 test result for the item-trait interaction (χ^2_{10}=13.6, *P*=.19, Bonferroni-adjusted *P*=.05/5=.01). The means and standard deviations of the fit residuals for items and persons were both within the acceptable limits. The single-item fit analyses for the remaining 5 categories were also satisfactory, with nonsignificant χ^2 test results (Table 4). The person separation index of this component was 0.68, and no differential item functioning was detected for any of the categories by gender, age, education level, or etiology.

Table 4. International Classification of Functioning, Disability, and Health categories retained after multiple rounds of Rasch analysis.

ICF[a] category		Location	Fit residual	χ^{2b}	*P* value	Transformation
Body functions and body structures						
1	Sleep functions (b134)	–1.450	–0.762	1.0	.61	4
2	Emotional functions (b152)	–1.281	–1.414	7.9	.02	3
3	Sensation of pain (b280)	0.308	1.047	2.4	.30	4
4	Blood vessel functions (b415)	2.034	0.395	3.7	.16	2
5	Increased blood pressure (b4200)	1.932	0.789	3.6	.16	3
6	Decreased blood pressure (b4201)	0.757	–1.106	1.6	.46	3
7	Immunological system functions (b435)	1.625	–0.922	1.8	.40	2
8	Respiration functions (b440)	1.767	–0.156	0.8	.68	3
9	Weight maintenance functions (b530)	0.729	0.523	2.4	.30	1
10	Sexual functions (b640)	–2.588	–2.086	5.5	.07	2
11	Procreation functions (b660)	–1.640	–0.862	1.6	.44	2
12	Mobility of joint functions (b710)	–0.895	–0.791	1.7	.43	2
13	Muscle power functions (b730)	–2.610	–0.216	3.5	.17	1
14	Muscle tone functions (b735)	–1.026	–0.586	2.7	.25	4
15	Structure of areas of skin (s810)	2.339	0.046	1.2	.54	4
Activities and participation[c]						
16	Changing basic body position (d410)	0.252	–1.543	5.6	.06	4
17	Moving around using equipment (d465)	–0.614	–0.274	2.2	.33	4
18	Washing oneself (d510)	–1.697	–1.594	5.2	.08	4
19	Caring for body parts (d520)	1.147	–0.637	0.8	.66	4
20	Regulating urination (d5300)	–0.866	3.353	8.5	.01	4
21	Regulating defecation (d5301)	–1.841	0.713	0.4	.80	4
22	Dressing (d540)	0.127	0.26	0.7	.72	4
23	Eating (d550) and drinking (d560)	3.493	–0.741	1.3	.52	4
Contextual factors[d]						
24	Assistive products and technology for personal indoor and outdoor mobility and transportation (e1201)	1.805	0.585	2.4	.30	2
25	Design, construction, and building products and technology of buildings for private use (e155)	–0.177	1.793	2.7	.26	2
26	Acceptance of life in a wheelchair/in bed	–0.237	–1.264	3.3	.19	2
27	Coping with everyday life	–1.127	1.305	1.9	.39	2
28	Adjustment to new body image	–0.263	–0.787	3.3	.19	2

[a]ICF: International Classification of Functioning, Disability, and Health.

[b]Goodness of fit of each category was tested using a χ^2 test with a Bonferroni-adjusted *P* value. All were <.01.

[c]Family relationships (d760) and socializing (d9205) were deleted because of poor fit.

[d]Knowledge about spinal cord injury was deleted because of poor fit.

Discussion

Principal Findings

The results of the Rasch analysis showed good fit to the Rasch model for the different components of the ICF set as implemented in the app after modification. Both overall and single-item fit were satisfactory. There was no differential item functioning for any of the ICF categories by gender, age, education level, or etiology. These results indicate the suitability of the app-based ICF set as an assessment tool for assessing the functioning of SCI patients.

The app-based ICF set is one of many forms of ICF-based electronic health records. Several previous studies have confirmed the role of ICF-based electronic health records in reflecting patient functioning and facilitating rehabilitation [6,34]. As a unified and standard language originally developed for multidisciplinary use, the original ICF was relatively easily understood by different disciplines and suitable for multidisciplinary teamwork [2]. However, some obstacles still existed, including the complexity of the terminology, lack of operationalization of the ICF qualifiers, and training overload for ICF users. The satisfactory internal construct validity of the set developed in the study is mainly due to the selection of suitable ICF categories and standardized assessment enabled by the app. In this study, the app applied many fewer ICF categories (31) than the comprehensive ICF core set (168 categories) and the brief core set (33 categories) for SCI patients issued by the WHO's ICF research branch [35]. Additionally, the categories were more specific because they were originally identified as good outcome indicators for SCI patients in China [20,26]. Each ICF category focuses on a specific area and is independent from the others, which helped to ensure the overall fit to the Rasch model based on the assessments' content.

Together's verbal prompts standardize assessment and give more consistent assessment results. The transformation guidelines operationalize the 5 ICF qualifiers simply and effectively. With the help of the app, the ICF qualifiers can automatically be matched to the initial clinical assessment results. No additional training on ICF terminology or qualifiers is needed. The process reduces the differences among assessors and makes presentation of the ICF data more convenient and intelligent.

Family relationships (d760), socializing (d9205), and knowledge about spinal cord injury were deleted. In a previous study, d760 and d9205 also exhibited poor model fit [13]. Family relationships are the basis for good functioning in the family, and socializing reflects a patient's social participation. Both of them are influenced by many factors such as age, severity of the injury, and financial considerations [36,37]. For an item to have perfect fit to the Rasch model, no other factors should influence a person's performance regarding the item except for the person's ability and the item's difficulty, which may explain why the two factors did not fit the Rasch model. The ICF does not classify personal factors because of the large social and cultural variance associated with them [1]. The personal factors assessed in this ICF set were identified by multiround expert surveys [26]. Knowledge about SCI is a broad concept covering many aspects such as injury outcomes, functional rehabilitation, and preventing complications, and it is influenced by multiple factors such as the patient's level of education, efficacy self-perceptions, and any health education they have received [38]. This may be why the item did not fit the Rasch model well. However, family relationships or support, social participation and the knowledge about SCI are important indicators for the outcomes of rehabilitation. To ensure the completeness of a patient's health records, it is suggested that they be retained as part of the basic health data collection but not as part of the assessment instrument. Further research should use professional measurements that reflect family relationships, socializing, and the knowledge level for SCI patients.

It is worth noting that, in contrast to the person separation index of the activities and participation component, the person separation index of the other two components was not ideal (0.5 and 0.68), which reflects the poor internal consistency of the two components. This may be related to the different measuring guidelines used. Four guidelines were used in the study to transform the input of the medical staff, based on patient information in the form of percentages, wording from patient reports, frequency with which a problem was observed, and scores on routine clinical instruments or standards. A previous study [29] found that the categories involving guideline 4 (which transforms scores of a routine clinical instrument or standards into the ICF qualifiers) had the best interrater reliability, and the categories involving guideline 2 (which transforms the wording from patient reports into the ICF qualifiers) had the lowest interrater reliability. In this study, all the activity and participation categories used the guideline with the highest reliability (guideline 4), while all the contextual factor categories used the guideline with the lowest reliability (guideline 2). Although the app automatically transforms the initial data into the ICF qualifiers, subjectivity and uncertainty still existed in the patients' initial reports and recall, which may explain the unsatisfactory person separation index of the two components.

Limitations

In interpreting these results, it is important to keep in mind that the relatively small sample may have influenced the representativeness of the results. Also, although the app was designed to assess SCI patients at home during transitional care via its remote follow-up function, the study data were collected in face-to-face interviews before the participants were discharged. Further validation with larger samples and remote assessment via the app's communication function are needed.

Conclusions

This study has confirmed the suitability of the Together app as an assessment tool. With relatively fewer but more specific ICF categories, it overcomes some of the limitations related to applying the ICF and makes assessment results more reliable and consistent. The app opens up a new way to use the ICF with SCI and electronic health records. In the future, the app could be used to capture information about the functioning of SCI patients at home remotely. Such assessment results would help to monitor patients' functional changes and differences, learn their needs, identify their problems, and provide evidences for further interventions if necessary.

Acknowledgments

The authors would like to express special thanks to Bing Xie and Haishan Pan, the engineers who provided technical support for the development of the app. This study was supported by China's National Natural Science Foundation (grant number 71603293)

and the Fundamental Research Funds for the Central Universities (grant number 20ykpy88). The sponsors had no role in the study's design; in the collection, analysis or interpretation of data; in writing of the report; or in the decision to submit the article for publication.

Authors' Contributions

MJ drafted the manuscript and assisted with managing the app. JT was responsible for conducting the study and revising the manuscript. SX, XH, and YW helped collect the data. TL was responsible for the management of the app. TY participated in the study's design. KL was responsible for the project design, implementation, quality control, app management, and manuscript revision. All the authors reviewed the submitted manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Together tutorial.
[MP4 File (MP4 Video), 14058 KB - mhealth_v8i11e20723_app1.mp4]

References

1. International Classification of Functioning, Disability and Health. Geneva: World Health Organization; 2001. URL: https://www.who.int/classifications/drafticfpracticalmanual2.pdf?ua=1 [accessed 2020-10-05]
2. Tempest S, Harries P, Kilbride C, De Souza L. Enhanced clarity and holism: the outcome of implementing the ICF with an acute stroke multidisciplinary team in England. Disabil Rehabil 2013;35(22):1921-1925 [FREE Full text] [doi: 10.3109/09638288.2013.766272] [Medline: 23530624]
3. Nilmart P, Vongsirinavarat M, Somprasong S, Apinonkul B. Development of an extensive assessment list for knee osteoarthritis based on the International Classification of Functioning, Disability and Health: a Delphi study. Int J Rehabil Res 2019 Sep;42(3):240-248. [doi: 10.1097/MRR.0000000000000361] [Medline: 31283550]
4. Kisser U, Adderson-Kisser C, Coenen M, Stier-Jarmer M, Becker S, Sabariego C, et al. The development of an ICF-based clinical guideline and screening tool for the standardized assessment and evaluation of functioning after head and neck cancer treatment. Eur Arch Otorhinolaryngol 2017 Feb;274(2):1035-1043. [doi: 10.1007/s00405-016-4317-6] [Medline: 27695934]
5. Vreeman DJ, Richoz C. Possibilities and implications of using the ICF and other vocabulary standards in electronic health records. Physiother Res Int 2015 Dec;20(4):210-219 [FREE Full text] [doi: 10.1002/pri.1559] [Medline: 23897840]
6. Stucki G, Bickenbach J. Functioning: the third health indicator in the health system and the key indicator for rehabilitation. Eur J Phys Rehabil Med 2017 Feb;53(1):134-138 [FREE Full text] [doi: 10.23736/S1973-9087.17.04565-8] [Medline: 28118696]
7. Bautz-Holter E, Sveen U, Cieza A, Geyh S, Røe C. Does the International Classification of Functioning, Disability and Health (ICF) core set for low back pain cover the patients' problems? A cross-sectional content-validity study with a Norwegian population. Eur J Phys Rehabil Med 2008 Dec;44(4):387-397 [FREE Full text] [Medline: 19002088]
8. Hilfiker R, Obrist S, Christen G, Lorenz T, Cieza A. The use of the comprehensive International Classification of Functioning, Disability and Health Core Set for low back pain in clinical practice: a reliability study. Physiother Res Int 2009 Sep;14(3):147-166. [doi: 10.1002/pri.436] [Medline: 19194959]
9. Starrost K, Geyh S, Trautwein A, Grunow J, Ceballos-Baumann A, Prosiegel M, et al. Interrater reliability of the extended ICF core set for stroke applied by physical therapists. Phys Ther 2008 Jul;88(7):841-851. [doi: 10.2522/ptj.20070211] [Medline: 18483130]
10. Uhlig T, Lillemo S, Moe RH, Stamm T, Cieza A, Boonen A, et al. Reliability of the ICF Core Set for rheumatoid arthritis. Ann Rheum Dis 2007 Aug;66(8):1078-1084 [FREE Full text] [doi: 10.1136/ard.2006.058693] [Medline: 17223659]
11. Cao R, Xu G, Ding X, Lin F, Li J. Study on the reliability and validity of international classification of function, disability and health brief core sets for stroke. Chin J Rehab Med 2011;26:715-719. [doi: 10.3969/j.issn.1001-1242.2011.08.005]
12. Gao Y, Yan T, You L, Li K. Developing operational items for the International Classification of Functioning, Disability and Health Rehabilitation Set: the experience from China. Int J Rehabil Res 2018 Mar;41(1):20-27. [doi: 10.1097/MRR.0000000000000254] [Medline: 28957982]
13. Li K, Yan T, You L, Xie S, Li Y, Tang J, et al. Psychometric properties of the International Classification of Functioning, Disability and Health set for spinal cord injury nursing based on Rasch analysis. Disabil Rehabil 2018 Feb;40(3):338-345. [doi: 10.1080/09638288.2016.1250169] [Medline: 27967253]
14. Barton AJ. The regulation of mobile health applications. BMC Med 2012;10:46 [FREE Full text] [doi: 10.1186/1741-7015-10-46] [Medline: 22569114]

15. Kaplan AL, Cohen ER, Zimlichman E. Improving patient engagement in self-measured blood pressure monitoring using a mobile health technology. Health Inf Sci Syst 2017 Dec;5(1):4 [FREE Full text] [doi: 10.1007/s13755-017-0026-9] [Medline: 29081974]
16. Semple JL, Sharpe S, Murnaghan ML, Theodoropoulos J, Metcalfe KA. Using a mobile app for monitoring post-operative quality of recovery of patients at home: a feasibility study. JMIR Mhealth Uhealth 2015;3(1):e18 [FREE Full text] [doi: 10.2196/mhealth.3929] [Medline: 25679749]
17. Wang Q, Zhao J, Huo X, Wu L, Yang L, Li J, et al. Effects of a home care mobile app on the outcomes of discharged patients with a stoma: a randomised controlled trial. J Clin Nurs 2018 Oct;27(19-20):3592-3602. [doi: 10.1111/jocn.14515] [Medline: 29775491]
18. Dahlberg K, Jaensson M, Nilsson U. "Let the patient decide": person-centered postoperative follow-up contacts, initiated via a phone app after day surgerycondary analysis of a randomized controlled trial. Int J Surg 2019 Jan;61:33-37 [FREE Full text] [doi: 10.1016/j.ijsu.2018.11.022] [Medline: 30503602]
19. De La Cruz Monroy MFI, Mosahebi A. The use of smartphone applications (apps) for enhancing communication with surgical patients: a systematic review of the literature. Surg Innov 2019 Apr;26(2):244-259. [doi: 10.1177/1553350618819517] [Medline: 30602332]
20. Liu T, Li K, Xie S, Wang Y, Tang J, He X, et al. Development of the APP for transitional care of people with spinal cord injury based on the ICF. Chinese J Rehab Med 2019;34:154-158. [doi: 10.3969/j.issn.1001-1242.2019.02.006]
21. Zhang N, Zhou M, Liu N, Liu X, Zhang Y. An investigation study on quality management of spinal cord injury rehabilitation in China 2016. Chinese J Rehab Med 2018;33:1137-1141. [doi: 10.3969/j.issn.1001-1242.2018.10.002]
22. Zeng X, Wang X, Xu J, Chen T, Xu J. Investigation on the rehabilitation of patients with amputees, spinal cord injuries and craniocerebral injuries after the earthquake in Shifang, Sichuan province. Chinese J Rehab Med 2010;25:771-774. [doi: 10.3969/j.issn.1001-1242.2010.08.014]
23. Chen B, Jiang C, Niu W. Research status of rehabilitation needs of patients with spinal cord injury in Wenchuan earthquake. Chinese J Hygiene Rescue 2017;3:72-74. [doi: 10.3877/cma.j.issn.2095-9133.2017.02.002]
24. Xie H, Yang Y, Wu A, Shen C, Hu L, Zhang J, et al. Status quo and development of hope houses for individuals with spinal cord injury in Shanghai from perspective of managers and organizers. Chinese J Rehab Theory Pract 2019;25:1221-1227. [doi: 10.3969/j.issn.1006-9771.2019.00.011]
25. SunYat-Sen University. Distance transitional care system for spinal cord injury patients. Copyright Protection Center of China Nov2018SR904705. 2018. URL: www.ccopyright.com [accessed 2020-10-09]
26. Li K, Xie S, Wang Y, Tang J, He X, Liu T, et al. Outcome indicators in the transitional care of people with spinal cord injury in China: a Delphi survey based on the International Classification of Functioning, Disability and Health. Disabil Rehabil 2019 Feb 06:1-9. [doi: 10.1080/09638288.2018.1528638] [Medline: 30724613]
27. Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. J Rehabil Med 2003 May;35(3):105-115 [FREE Full text] [doi: 10.1080/16501970310010448] [Medline: 12809192]
28. Schuld C, Franz S, Brüggemann K, Heutehaus L, Weidner N, Kirshblum SC, EMSCI study group. International standards for neurological classification of spinal cord injury: impact of the revised worksheet (revision 02/13) on classification performance. J Spinal Cord Med 2016 Sep;39(5):504-512 [FREE Full text] [doi: 10.1080/10790268.2016.1180831] [Medline: 27301061]
29. Li K, Yan T, You L, Xie S, Li Y, Tang J, et al. The inter-rater reliability of the International Classification of Functioning, Disability and Health set for spinal cord injury nursing. Int J Rehabil Res 2016 Sep;39(3):240-248. [doi: 10.1097/MRR.0000000000000176] [Medline: 27171607]
30. Grill E, Stucki G. Scales could be developed based on simple clinical ratings of International Classification of Functioning, Disability and Health Core Set categories. J Clin Epidemiol 2009 Sep;62(9):891-898. [doi: 10.1016/j.jclinepi.2008.02.003] [Medline: 18619794]
31. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ 1995 Jan 21;310(6973):170 [FREE Full text] [doi: 10.1136/bmj.310.6973.170] [Medline: 7833759]
32. Cieza A, Hilfiker R, Boonen A, Chatterji S, Kostanjsek N, Ustün BT, et al. Items from patient-oriented instruments can be integrated into interval scales to operationalize categories of the International Classification of Functioning, Disability and Health. J Clin Epidemiol 2009 Sep;62(9):912-921. [doi: 10.1016/j.jclinepi.2008.04.011] [Medline: 19541452]
33. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum 2007 Dec 15;57(8):1358-1362 [FREE Full text] [doi: 10.1002/art.23108] [Medline: 18050173]
34. Maritz R, Aronsky D, Prodinger B. The International Classification of Functioning, Disability and Health (ICF) in electronic health records. a systematic literature review. Appl Clin Inform 2017 Sep 20;8(3):964-980. [doi: 10.4338/ACI-2017050078] [Medline: 28933506]
35. ICF-based documentation tool: ICF Core Sets in clinical practice. Geneva: World Health Organization; 2012. URL: https://www.icf-core-sets.org/ [accessed 2020-10-05]
36. Noreau L, Fougeyrollas P. Long-term consequences of spinal cord injury on social participation: the occurrence of handicap situations. Disabil Rehabil 2000 Mar 10;22(4):170-180. [doi: 10.1080/096382800296863] [Medline: 10798305]

37. Kennedy P, Lude P, Taylor N. Quality of life, social participation, appraisals and coping post spinal cord injury: a review of four community samples. Spinal Cord 2006 Feb;44(2):95-105. [doi: 10.1038/sj.sc.3101787] [Medline: 16130026]
38. Zhang W, Li S, Li X, Chen C, Tang L, Hu J, et al. [Currency survey and influencing factors analysis on the knowledge and behavior related to brucellosis among occupational workers in Jianyang City]. Sichuan Da Xue Xue Bao Yi Xue Ban 2019 Jul;50(4):551-555. [Medline: 31642234]

Abbreviations

ICF: International Classification of Functioning, Disability, and Health
NRS: numerical rating scale
SCI: spinal cord injury
WHO: World Health Organization

Edited by G Eysenbach; submitted 27.05.20; peer-reviewed by C Jacob, L Grepo; comments to author 15.08.20; revised version received 26.09.20; accepted 29.09.20; published 11.11.20.

Please cite as:
Jia M, Tang J, Xie S, He X, Wang Y, Liu T, Yan T, Li K
Using a Mobile App-Based International Classification of Functioning, Disability, and Health Set to Assess the Functioning of Spinal Cord Injury Patients: Rasch Analysis
JMIR Mhealth Uhealth 2020;8(11):e20723
URL: https://mhealth.jmir.org/2020/11/e20723
doi:10.2196/20723
PMID:33174860

Original Paper

Patients and Medical Staff Attitudes Toward the Future Inclusion of eHealth in Tuberculosis Management: Perspectives From Six Countries Evaluated using a Qualitative Framework

Ioana Margineanu[1,2], MD; Christina Louka[3], MD; Maria Vincenti-Gonzalez[4], PhD; Antonia Morita Iswari Saktiawati[3,5], PhD; Johannes Schierle[6], BA; Kabiru Mohammed Abass[3,7], MD; Onno Akkerman[8], MD, PhD; Jan-Willem Alffenaar[1,9,10,11], Prof Dr, PharmD; Adelita V Ranchor[12], Prof Dr; Ymkje Stienstra[3], MD, Prof Dr

[1]Department of Clinical Pharmacy and Pharmacology, University Medical Centrum Groningen, University of Groningen, Groningen, Netherlands

[2]Pneumology Hospital Iasi, Iasi, Romania

[3]Department of Internal Medicine, University Medical Centrum Groningen, University of Groningen, Groningen, Netherlands

[4]Department of Medical Microbiology and Infection Prevention, University Medical Centrum Groningen, University of Groningen, Groningen, Netherlands

[5]Department of Internal Medicine, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia

[6]University Medical Centrum Groningen, University of Groningen, Groningen, Netherlands

[7]Agogo Presbyterian Hospital, Agogo, Ghana

[8]Department of Pulmonary Diseases and Tuberculosis, Tuberculosis Centrum Beatrixoord, University Medical Centrum Groningen, University of Groningen, Haren, Netherlands

[9]University of Sydney, Camperdown, Australia

[10]Westmead Hospital, Sydney, Australia

[11]Marie Bashir Institute of Infectious Diseases, University of Sydney, Sydney, Australia

[12]Health Psychology Section, University Medical Centrum Groningen, University of Groningen, Groningen, Netherlands

Corresponding Author:
Ioana Margineanu, MD
Department of Clinical Pharmacy and Pharmacology
University Medical Centrum Groningen
University of Groningen
Hanzeplein 1
Groningen
Netherlands
Phone: 1 61 496 3518
Email: ismargineanu@gmail.com

Abstract

Background: Digitally delivering healthcare services is very attractive for tuberculosis (TB) management as this disease has a complex diagnosis and lengthy management and involves multiple medical and nonmedical specialists. Especially in low- and middle-income countries, eHealth could potentially offer cost-effective solutions to bridge financial, social, time, and distance challenges.

Objective: The goal of the research is to understand what would make eHealth globally applicable and gain insight into different TB situations, opportunities, and challenges.

Methods: We performed focus group interviews with TB experts and patients from 6 different countries on 4 different continents. The focus group interviews followed the theory of planned behavior framework to offer structured recommendations for a versatile eHealth solution. The focus group interviews were preceded by a general demographic and technology use questionnaire. Questionnaire results were analyzed using basic statistics in Excel (Microsoft Corporation). Focus group interview data were analyzed using ATLAS.ti 8 (ATLAS.ti Scientific Software Development GmbH) by assigning codes to quotations and grouping codes into the 5 domains within the framework.

Results: A total of 29 patients and 32 medical staff members were included in our study. All medical staff had used the internet, whereas 31% (9/61) of patients had never been online. The codes with the most quotations were information in relation to eHealth

(144 quotations) and communication (67 quotations). The consensus among all participants from all countries is that there are important communication and information gaps that could be bridged by an eHealth app. Participants from different countries also highlighted different challenges, such as a majority of asylum-seeker patients or lack of infrastructure that could be addressed with an eHealth app.

Conclusions: Within the 6 countries interviewed, there is high enthusiasm toward eHealth in TB. A potential app could first target information and communication gaps in TB, with additional modules aimed at setting-specific challenges.

(JMIR Mhealth Uhealth 2020;8(11):e18156) doi:10.2196/18156

KEYWORDS

eHealth; tuberculosis; policy; clinical; patient; perspective

Introduction

The continuous growth of the internet and availability of smart technologies have modified many aspects of life, including health care delivery. With more than half the world's population online [1], delivering health-related services digitally has never been more appealing or accessible, with the eHealth market expected growth estimated at 22% by 2024 [2].

Tuberculosis (TB) is one the deadliest infectious diseases worldwide, with an estimated 1.5 million deaths yearly [3]. The more resistant forms, like multidrug-resistant TB and extensively drug-resistant TB, pose a new threat, with treatment success rates between 50% to 60% globally. TB diagnosis is complex, its treatment is lengthy, and it requires close collaboration of different medical and nonmedical experts and patients to ensure TB management is adequately performed, especially in challenging settings such as patients living in remote locations or constrained by socioeconomic factors.

Studies piloting various types of eHealth technologies conducted around the world have evaluated multiple areas where eHealth could aid in TB management. From improving communication among medical staff [4] and patients [5] to reducing costs [6] and improving treatment indicators, especially in situations where patients were traveling [7] or in hard to reach regions [8], eHealth seems to be a promising field of research and a useful, cost-efficient, and acceptable improvement for TB management.

On the other hand, multiple studies and reports, including a consumer report from the European Union [9], have observed that there are certain difficulties when implementing eHealth, such as lack of acceptance, unfavorable regulations, and insufficient funding [10]. The progress of eHealth in lower income countries is limited by lack of know-how, funding, technology, and communications ability [11].

Taking into consideration the particular nature of TB and population differences together with the potential benefits and challenges of eHealth, tailored research would aid in creating useful, tailor-made eHealth apps. This idea is recommended in multiple studies, including a recent review of smartphone apps [12]. Market research, defined as an "organized effort to gather information about target markets" [13], is an important component of any business strategy aiming to create a new product. In order to make sure a product is useful, acceptable, and qualitative, market research is performed to understand the needs and preferences of the market. One of the frameworks used to conduct market research is the theory of planned behavior [14], which links psychological intent to three determinants: attitude toward behavior, subjective norm, and perceived behavioral control. This theory attempts to explain behavioral intentions and has been previously used in numerous studies to investigate and explain intention and possible adoption rates of new interventions [15,16].

This study aims to identify potential user perceptions about eHealth use in TB management. We used the theory of planned behavior framework to interview TB experts and patients in 6 diverse countries on 4 different continents to form recommendations for a well-received, usable, comprehensive, and efficient TB eHealth app.

Methods

Participants

Adult (over aged 18 years) participants from Romania, Greece, Netherlands, Indonesia, Ghana, and Venezuela were approached in the collaborating clinics and invited to participate. Two participant groups per country were interviewed, one comprising actual or former TB patients and one of the TB medical staff experts. Experts had to have worked in a TB clinic on a daily basis for at least 3 years. Participants were recruited through purposive sampling by investigators visiting different medical facilities specializing in TB care. The target size was 2 to 6 participants per focus group as all participants were highly involved with the topic of TB and discussion in larger groups was not recommended [17].

Study Design

The study was performed in two parts. The first phase consisted of a short questionnaire to collect demographic data and basic internet and mobile use statistics by asking the experience, in years and number of hours per day, spent using the internet and smartphones and a numbered scale on which participants ordered activities performed in order of frequency (with 5 being most frequent and 1 least frequent). Activities were defined as communication (eg, WhatsApp, Facebook Messenger), social media (eg, Facebook, Instagram), utilities (eg, banking, weather), work, games, and health/medical.

The second part consisted of semistructured focus group interviews with questions as conversation starters. The researchers allowed the interviewees to have a conversation around a specific question and asked follow-up questions for clarification. Questions were designed to repeatedly ask the

same subject in different ways and at different time points during the interview to achieve data saturation irrespective of time needed. In order to adhere to qualitative reporting standards, the consolidated criteria for reporting qualitative research (COREQ) checklist was used [18].

In order to minimize the risk of bias, the researcher conducting the interviews had no previous history with the patients or medical staff involved and conducted the interviews in a private room, patients separated from medical staff. Interviews were conducted in the participants' native tongues or in a language they felt comfortable in, transcribed verbatim, and translated by a native speaker with proficiency in English.

Framework

In order to develop a globally acceptable app that can be used in different settings, interviews targeted diverse countries with different cultures, socioeconomic statuses, and TB populations. Thus, 6 countries, Romania, Ghana, Indonesia, Greece, the Netherlands, and Venezuela, were chosen to have a mix of geography, socioeconomic statuses, health care systems, and TB profiles (Multimedia Appendix 1).

The focus groups followed a semistructured framework based on the theory of planned behavior (Textbox 1). This psychological theory proposes that intention—in our case, use of eHealth—has a number of determinants. The original work describes three main determinants, and these were used in order not only to stay true to the framework, but also to guide interviews in a simple, efficient manner. The first is the attitude toward the behavior, defined by the strength of the attitude and the evaluation of the outcome, favorable or unfavorable. The second is the subjective norm, or beliefs about the normative expectations of others: perceived social pressure to perform or preclude from the behavior. The last is the perceived ease or difficulty of performing the behavior and is based on past experience as well as anticipated factors that might facilitate or impede a specific behavior [14]. The last domain, preferred features, was added to further stratify user preferences.

Textbox 1. Interview structure based on the theory of planned behavior.

Attitude toward behavior:

- Q1: Which problems and challenges in tuberculosis management could be solvable by an eHealth solution?
- Q1 follow-up: If you had to identify priorities, which would be the biggest challenge?

Subjective norm:

- Q2: How do you think the medical staff and patients here would react to the implementation of an eHealth solution?
- Q2 follow-up: What do you think will be the biggest problem or motivator to accept and use eHealth?

Perceived ease or difficulty of adopting behavior:

- Q3: Do you use technology regularly to assist with your work or patient life (for admitted patients)?
- Q3 follow-up A: Do you think the implementation of various software solutions has made your life (work/patient) easier or harder?
- Q3 follow-up B: Why?
- Q3 follow-up C: What could be done to a new eHealth app to make it really useful?
- Q4: Do you think implementing eHealth in your daily lives will be easy or difficult?
- Q4 follow-up A: Why?
- Q4 follow-up B: What would make it easier or harder to implement?

Preferred features:

- Q5: Name 5 features or things you would likely use most in an eHealth app.
- Q5 follow-up: Which one do you feel you need most? Which process of tuberculosis management would be most suitable to be streamlined through eHealth?

Ethics

The study was approved by the ethics committee of the initiating institute, the University Medical Centrum Groningen (METc 2017/448), and the medical facilities of each country participating in the study. All focus group participants signed informed consent expressing their volition to participate in the study.

Data Collection and Analysis

A thematic approach using ATLAS.ti 8 (ATLAS.ti Scientific Software Development GmbH) was used to analyze the transcripts by one investigator (IM); this was reviewed by the supervisory authors. After major themes were identified in response to the questions asked and the theory of planned behavior framework, the transcriptions were indexed using topic coding (Textbox 2). Codes were assigned to phrases addressing an issue in the positive or negative (eg, "Maybe if we use an app, [the process] would be made simpler" versus "[eHealth] would give us double the job") and were assigned whenever a

quote was repeated by other participants in the focus group but not by the same participant. For clarification purposes, see definitions in Multimedia Appendix 2. All of the coding was performed by two independent authors (IM, CL). Conflicts were resolved by discussions between the coders with the aid of one of the supervisors (YS).

Descriptive statistics were used to present quantitative results. Medians and interquartile ranges were used to present internet and mobile experience. Codes were reported as a total and per domain, and themes were reported as proportions of the total.

Textbox 2. Coding.

Attitude toward behavior:

- communication – eHealth
- eHealth – speed
- eHealth – would/wouldn't help in tuberculosis
- information/education – eHealth
- my own experience with tech

Subjective norm:

- communication – status quo
- how will others react to eHealth in tuberculosis
- information – status quo
- local environment
- others' experiences with tech
- stigma/isolation

Perceived ease/difficulty of adopting behavior:

- Community building/testimonials
- I can/I can't implement eHealth
- Localization
- Money/devices
- Privacy/confidentiality
- Repeatability
- User interface/simplicity
- Training
- Videos

Preferred features:

- Screening/prevention
- Diagnosis
- Treatment
- Reminders
- Gamification
- Media preference
- Anything else

Results

General Questionnaire Results

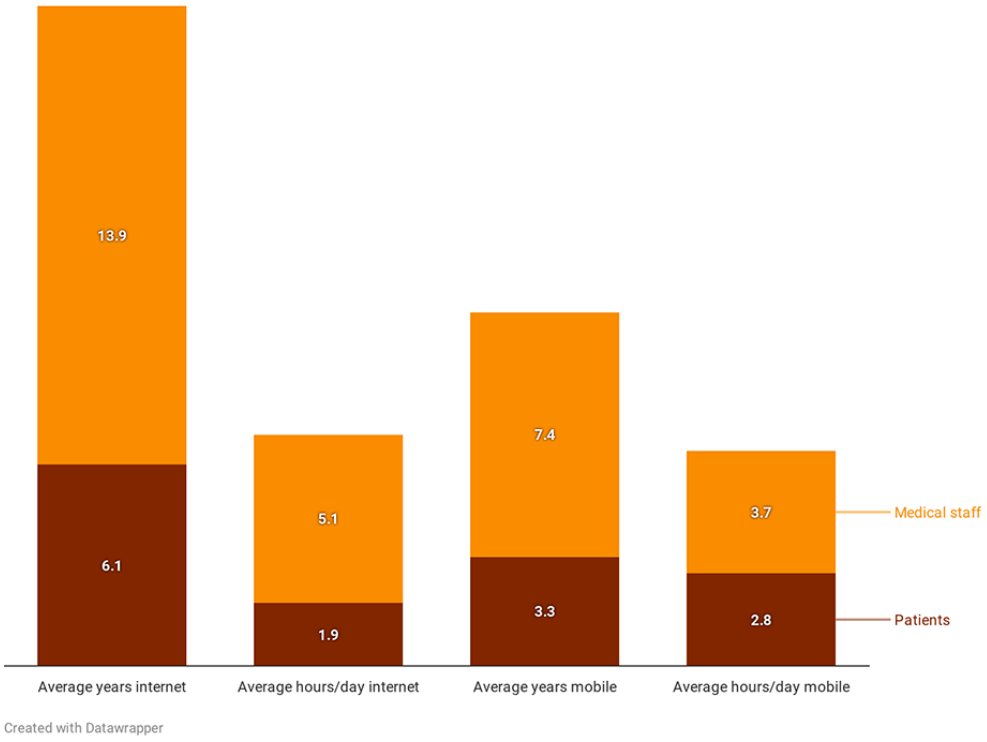
A total of 29 patients and 32 medical staff members responded to the questionnaire and participated in the focus group. Four patients were former TB patients and had finished their treatment within 6 months of the interview, with the majority being current patients. Ages ranged between 23 and 63 years for patients and 30 and 60 years for medical staff. Gender distribution was mostly male (20/29, 68%) among patients and mostly female (26/32, 81%) among medical staff; 65% (19/29) of patients and 68% (22/32) of medical staff lived in urban environments.

Education levels were lower in patients groups than in the medical staff group: among patients, the highest level of education was a bachelor's degree, and among medical staff, the highest level was doctorate (PhD).

General internet and mobile internet use was almost half within the patient group compared with the medical staff (Figure 1): average years of experience using the internet was 6.1 (range 1-16) person-years versus 13.9 (range 5-35) person-years. All medical staff had used the internet, whereas 31% (9/29) of patients had never been online. Concerning mobile phone use, staff median experience was 7.4 (range 0.3-15) person-years versus 3.4 (range 0-10) person-years for patients; 44% (13/29) of patients had never used mobile data and 34% (10/29) had never used a smartphone. Of all patient groups, the least tech savvy were the Romanians and the most were from Venezuela. From the medical staff groups, participants from Ghana had the least experience with the internet and mobile, and participants from the Netherlands had the most. Internet and mobile use in hours per day were lower for people from the low- and middle-income countries than the high-income countries, but this correlation was observed only in the patient groups and not in the expert groups.

The 20 patients who used the internet mostly used it for communication (median 4/6), followed by social media (median 4/6), utilities (median 3/6), medical/health (median 2/6), work (median 1/6), and games (median 1/6). For smartphones, the 18 patients ordered the categories the same, but with more patients using Messenger (median 4.5/6) and fewer using for social media and utilities (median 3/6). The 32 staff members used the internet mostly for communication and work (median 5/6); games were the least used (median 1/5). Medical staff use smartphones also mostly for communication (median 5/6), followed by social media (median 4/6), and work (4/6).

Figure 1. Internet and mobile experience for the two participant groups.



Focus Group Interviews

The final yield was 13 focus group interviews, two per country, with 2 focus groups conducted with Dutch medical staff as the TB clinic and at the Municipal Health Center Groningen (a type of outpatient clinic) are geographically separated. The duration of interviews ranged between 25 and 45 minutes.

By domain, attitude toward behavior contained the most quotes (339), with preferred features containing the least (161). By far the most discussed topic was the potential of eHealth to improve information and/or education concerning TB among patients, the public, and even medical staff (144 quotes). The countries that mentioned this aspect the most were Venezuela (55 quotes) and the Netherlands (31 quotes).

A summary of the most frequent themes/codes and example quotations can be found in Multimedia Appendix 2.

Thematic and Code Analysis

The attitude toward behavior domain contained 5 codes, with a total yield of 339 quotations out of the total 911 quotations (37.2%), the most of all domains. Even though experience varied between patients and medical staff, both groups explained that their outlook concerning implementing eHealth in TB was overwhelmingly positive (292/339, 86.1%), with the most expected impact in the information and communication fields.

> *Education [would be the first priority when implementing eHealth].* [Medical staff, Romania]

> *Actually there is already a program for TB managing. It has been running actually, but not everyone used it [referring to an electronic database for medical staff].* [Medical staff, Indonesia]
>
> *Maybe if using an app, would make it simpler.* [Medical staff, Indonesia]

Participants also offered examples of specific ways in which eHealth could improve TB management.

> *It's important to really explain that the disease is treatable and to know more about this disease and to say that people can survive.* [Patient, Venezuela]

Participants felt that speed, convenience, flexibility, and the ability to communicate and obtain information over long distances were very important factors contributing to a positive outcome.

> *And also the plain fact that they wouldn't have to come to the hospital could be a motive [to adopt an eHealth app].* [Medical staff, Greece]

Although the expected outcome is mostly positive, some challenges have been identified, such as a need for personal contact. Interestingly, within the same focus group, one person believed an app would enhance contact and others did not.

> *The fault is the lack of contact in DOT and an app could do that.* [Medical staff, the Netherlands]
>
> *Yes, but nothing can replace the live contact.* [Medical staff, the Netherlands]

Furthermore, patients mentioned a need for trustworthy sources of information, and some medical staff were concerned about patients intruding on their personal time if the app were to be on their personal phones or of doubling the workload (traditional system and eHealth).

The subjective norm domain contained 6 codes with 22.8% (208/911) of the total quotations. All groups felt there is a certain degree of pressure to perform the behavior as they felt the status quo should be improved, and they thought eHealth could be a useful tool for such improvement. Participants described gaps especially in communication and information access that would encourage them to use a trusted eHealth app.

> *There needs to be more explanations about what you can do, more explanations about the multiple versions of TB. We don't know about that you can have it in your bones, people don't know about adverse reactions, we don't know about these things.* [Patient, the Netherlands]
>
> *Sometimes it [the internet search] will not give you a straight answer. Sometimes it will lead you to further look.* [Patient, Ghana]

Concerning the particular use of eHealth in TB and the perceived subjective norm concerning the general public, patients spoke about stigma and how eHealth, through education, could solve such problems.

> *For example, I feel a lot of rejection every time; people changed a lot when they found I am TB positive, they don't even say hello directly to me, but from afar and this rejection is due to the lack of knowledge, education.* [Patient, Venezuela]

Participants believed either that eHealth would be easy to implement or that it is possible but that ease of implementation will depend on certain factors, most frequently technology savviness.

> *If I could use technology, I would use it. But now I am afraid I would make mistakes whilst using it because I would get confused.* [Patient, Greece]
>
> *Our TB patients are not there yet. We don't really have university professors with TB.* [Medical staff, Romania]

Participants also mentioned lack of resources as possible impediments to eHealth implementation. Lower income countries tended to mention more often a lack of physical resources, such as electricity, network coverage, or hardware (computers, phones), whereas higher income countries were worried about the lack of human resources needed to manage a potential extra burden of eHealth.

> *They [the patients] get their phones stolen, or all sorts of cables can be robbed.* [Medical staff, Venezuela]
>
> *The [clinic] does not have the manpower to see everyone, they [clinic staff] go [to see patients] maybe once every two months.* [Medical staff, the Netherlands]

Perceived ease or difficulty toward behavior contained 9 codes and 22.2% (203/911) of the total quotations. A minority of participants expressed they didn't think they could use a new app, quoting reasons such as illiteracy or lack of time. Most either said they would have no problems integrating a new app within their digital routines or that they would require certain facilitators that would promote adoption (eg, have a simple interface, contain a training module).

> *We are always open to new technology which could help us improve even more.* [Medical staff, Greece]
>
> *[Asked how they would respond to new app] Easy, easy [raised voice, altogether].* [Patients, Venezuela]
>
> *And if you have too many apps, like we already do, to have another one, it could be time consuming. We wouldn't want to be overloaded.* [Medical staff, the Netherlands]

Participants identified steps that are mandatory to perform in order to have an easy-to-use app, such as privacy or localizing the app.

> *Privacy, that's what I would consider first. There shouldn't be any breaches because otherwise it wouldn't succeed so no one should have access to data.* [Patient, Greece]

Facilitators concerning implementation, such as financial incentives in the form of extra staff or equipment or offering the app for free were also identified, especially since some participants expressly mentioned the financial difficulties some patients are in.

> *I believe if the app could be in every language and the devices would be provided people would be very grateful, so it won't also be useful, but they would be more compliant.* [Medical staff, Greece]
>
> *Then the cost. It should be free.* [Patient, Ghana]

Some participants expressed a desire to be trained in how to use the app or for a demo module to be presented as they believe this would facilitate quick adoption.

> *Some of us will need someone to teach us.* [Patient, Romania]

Concerning the app itself, participants believed ease of use would be furthered by a simple and friendly user interface, by adding a community-building module (either for the patients or for medical staff), adding a training module, and using videos to transmit information.

> *It should be simple, with not many things, because it would demotivate me.* [Medical staff, the Netherlands]
>
> *It can also even have people who are also being successfully treated, people like that communicating, all that being put into the app so that when the person goes, person knows from the beginning how the treatment goes, and after treatment what to expect.* [Medical staff, Ghana]

Preferred features contained 7 codes with 17.6% (161/911) of the total quotations. Concerning information media, most participants expressed a preference for video, followed by images, text, and a gamification component, such as quizzes, to enhance the user experience.

> *I think film would be nice. When you read you can put it away.* [Medical staff, the Netherlands]
>
> *Or maybe a countdown, how far you are and how much you have left and then patients know what to eat. For them to see how easy it is and to stimulate them to continue.* [Medical staff, the Netherlands]

Participants would rather use the app for treatment (38/61) than for diagnosis (17/61) or prevention (19/61), and they would want a notification or reminding system implemented to encourage treatment adherence and follow-up.

> *Patients can use an application to schedule visits to health facilities, start treatment, when should do sputum test, do monthly check up. Things that they usually do here, but it will be paperless/electronically. Then all these [digital] notes can be brought everywhere, I mean paper note can be lost, can be damaged.* [Medical staff, Indonesia]
>
> *Through this media/app patients can be educated and we can reinforce prevention.* [Medical staff, Venezuela]

Discussion

Principal Findings

Using a health behavior framework for market research [19], this study explored attitudes toward eHealth implementation in TB. Semistructured focus group interviews were performed worldwide with medical staff and patients to better understand key motivators, challenges, facilitators, and user preferences for implementing new eHealth solutions. A number of important insights have been gathered, as has a prioritization of features to be implemented, which can be used when planning new eHealth apps for TB. Overall, both patients and TB experts have expressed enthusiasm at the potential of eHealth, with an overwhelming consensus that the first domains where it could be useful are information and communication.

The attitude toward eHealth domain contained the most codes. We interpreted this result in the context of participants having already formed an opinion on this subject and welcomed the opportunity to discuss it in depth. The domain presents encouraging results, with 67% (229/339) of codes expressing a positive expectation about eHealth capabilities. Participants overwhelmingly felt there is a lack of knowledge about TB among patients, the general public, and even among medical practitioners. These findings are mirrored by a systematic review that concluded there is a lack of knowledge among medical practitioners concerning national or internal TB guidelines within 14 non-European countries [20] and by multiple studies identifying knowledge gaps among patients [21-23]. Furthermore, a recent review highlighted that many apps offer inaccurate information [24]. Participants expect that an eHealth solution would be used to educate and provide accurate, secure, and friendly information.

The subjective norm describes the most important pressure to adopt eHealth as the lack of information and communication, felt across the board. Participants described a lack of clear, open communication channels both between patients and medical staff and between different specialties involved in TB management. Multiple studies have linked lack of patient-medical staff relationships to poorer outcomes in TB [25-27], concluded that communication methods should be tailored [28], or called for improving collaborations between medical staff involved in TB care [29]. Participants in our study would not only welcome online TB-related communities but also believe that communication could be improved through an app and that, in itself, would improve TB management and the stigma felt at the moment. On the other hand, a challenge identified was the need for human, personal contact, identified especially by a minority of Dutch participants (2).

Concerning perceived ease or difficulty of use, most participants felt they could implement a new app easily, although some participants mentioned a fear that a new app would be time consuming.

One important factor that could influence adoption was identified as technology savviness, linked to age and experience with use; however, some participants felt that training could bridge this gap. Indeed, one study using video directly observed therapy noted that "Older participants in particular enjoyed learning to use a smartphone" [30]. A facilitator mentioned by some participants is localization, translating the app in the local language. A minority of participants quoted illiteracy as a barrier to regular app use.

On a local level, interviewees from countries with a resource paucity, such as Venezuela or Greece, expressed a need for

extra human or physical resources. Most participants agreed that an app should be free. Interestingly, concerning technology accessibility, opinions varied widely, from "even if they are illiterate they use the internet, google, even if they can't write their own signature, they can go online" to "they probably have 1 smartphone per family and they don't use the internet all the time."

From a development perspective, the only truly mandatory feature to be implemented would be privacy/security, as this was a concern expressed in multiple interviews, both by experts and patients. A recent systematic review performed by the authors (unpublished) highlighted that within 7 studies that quantified this aspect, there were zero privacy breaches for a pooled 71 patients.

Participants mentioned other features, such as a treatment module with asynchronous video therapy, reminders, and a diagnosis module with an emphasis on self-diagnosing education in order to hasten hospital visits. Participants also expressed a desire to have a gamification component and believed video would best facilitate app adoption.

Textbox 3 summarizes recommendations for developing a new eHealth app for TB.

Textbox 3. Recommendations for developing a new eHeath app for tuberculosis.

General:

- Target the app for education, followed by communication, treatment, prevention, and diagnosis
- Identify a multidisciplinary team that can create a universally usable app, but recruit local members to advise, tailor-make, and translate
- Create a concept revolving around modules or build separate apps (eg, for medical staff and for patients)

App-related:

- Make sure the app is private and secure
- Focus on a simple and friendly user interface with a preference for images/icons over text
- Create training and educational videos
- Offer the app for free

Make sure the app is private and secure:

- Use an institutional server if your institution already has taken security measures for its existing patients
- Use an external secure server for additional privacy
- Respect local and regional privacy laws (eg, General Data Protection Regulation in the European Union)

Implementation:

- Find local tech leaders who can learn the app quickly and facilitate dissemination and further education
- Preferably, offer devices for at least the most disadvantaged members of the target populace
- If designing a medical staff app, discuss with administrators so that the app decreases the workload and doesn't add to it
- Have an active tech support

Modules:

- Education module, preferably with a gamification component
- Communication components:
  - medical staff (results, expert forum)
  - patients (testimonials, community)
  - medical staff and patients (side effects, questions, and appointments)
- Treatment module: calendar, video/directly observed therapy, reminders

Comparison With Prior Work

There are few studies exploring user attitudes for eHealth in TB. A study from Mozambique exploring text messaging found, as our study did, that messaging should be used for reminders and motivational texts in order to increase retention and that the main obstacle would be privacy assurance [31]. Another study involving focus group interviews with medical staff for a chronic obstructive pulmonary disease telerehabilitation app found that education and skill training are highly essential to support successful implementation [32]. A study from Saudi Arabia showed that perceived usefulness and perceived ease are significant factors to performing the behavior, results corroborated by our study [33]. Furthermore, a recent systematic review (unpublished) performed by the main authors revealed that already implemented eHealth apps focused rarely on

education, despite it being one of the two major needs felt by the participants within the focus groups.

This study offers a more diverse perspective on eHealth use in TB by conducting interviews in 6 countries on 4 different continents to gain a more global perspective for a potential app that could be universally applicable.

Limitations

Purposive sampling was used, which might have selected participants already more open to new technology. Furthermore, individual interviews might have elicited different results as they have less risk of bias and are more in-depth. The theory of planned behavior is a useful tool to gauge decision making with the caveat that it does not take in account socioeconomic, religious, and gender-based factors.

Strengths

This study was conducted in 6 different settings, offering a better understanding of what populations across the globe might decide concerning adopting eHealth in TB management. TB patients and TB experts were approached, thus covering the potential user base for future eHealth solutions. The theory of planned behavior is a simple, elegant way to conduct focus group interviews and understand decision-making processes. Two independent authors coded the interviews, thus limiting bias.

Conclusion

This study used focus group interviews performed in 6 countries in order to gauge perceptions about eHealth use in TB management and draw recommendations for further implementation. Participants in all 6 countries are enthusiastic about eHealth, and most users expect a potential app to be helpful. There is a global need to improve information access and communication, and participants feel that eHealth could help bridge this gap. Most themes resounded in all countries interviewed, but certain particularities, such as a large proportion of asylum seekers or lack of infrastructure or training, should be addressed when trying to implement eHealth in specific settings. Despite individual preferences, the global sentiment is that eHealth is a promising field of research that will be well received with the potential to enhance multiple aspects of TB care, with an emphasis on the need to communicate and educate.

Acknowledgments

The study team would like to thank the medical facilities where the studies took place: Iasi Pulmonology Hospital, Beatrixoord Tuberculosis Centrum, Gemeentelijke Gezondheidsdienst Groningen, Presbyterian Hospital Agogo, Thoracic Diseases General Hospital Sotiria, and Sardjito Hospital. This study was made within a doctoral project funded by the European Union Horizon 2020 Research and Innovation Program under the Marie-Skłodowska Curie grant agreement 713660. The funding source had no impact on any decision making regarding this paper.

Authors' Contributions

IM and YS initiated and designed the study. IM coordinated the interviews, translated the Romanian interview, translated (with the help of a translator) the Netherlands interview, participated in interview transcription, coded the interviews (independently), wrote the initial draft of the manuscript, and created the figures. CL conducted and translated the interview in Greece, participated in interview transcription, coded the interviews (independently), and reviewed the manuscript and figures. MVG conducted and translated the interview in Venezuela, participated in interview transcription, and reviewed the manuscript and figures. MS conducted and translated the interview in Indonesia, participated in interview transcription, and reviewed the manuscript and figures. JS and KMA conducted and translated the interview in Ghana, participated in interview transcription, and reviewed the manuscript and figures. OA, JWA, AVR, and YS supervised all stages of the study process and reviewed the manuscript and figures.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Country profiles.
[DOCX File , 13 KB - mhealth_v8i11e18156_app1.docx]

Multimedia Appendix 2
Definitions.
[DOCX File , 13 KB - mhealth_v8i11e18156_app2.docx]

Multimedia Appendix 3
Themes and codes ordered by frequency.
[DOCX File , 22 KB - mhealth_v8i11e18156_app3.docx]

References

1. Scott F, Meagher S. ASUM press release. Australas J Ultrasound Med 2011 Nov;14(4):2 [FREE Full text] [doi: 10.1002/j.2205-0140.2011.tb00128.x] [Medline: 28191121]
2. Global eHealth market size, share, trends and growth analysis report – segmented by type (telemedicine, consumer, information system, HER, ePrescribing, health information, mHealth, health management and clinical decision support), services, end user and region - industry forecast | 2020 to 2025. 2020 Feb. URL: https://www.marketdataforecast.com/market-reports/e-Health-market [accessed 2020-10-20]
3. Global tuberculosis report 2019. Geneva: World Health Organization URL: https://apps.who.int/iris/bitstream/handle/10665/329368/9789241565714-eng.pdf?ua=1 [accessed 2020-10-07]
4. Blaya JA, Shin SS, Yagui M, Contreras C, Cegielski P, Yale G, et al. Reducing communication delays and improving quality of care with a tuberculosis laboratory information system in resource poor environments: a cluster randomized controlled trial. PLoS One 2014;9(4):e90110 [FREE Full text] [doi: 10.1371/journal.pone.0090110] [Medline: 24721980]
5. Kunawararak P, Pongpanich S, Chantawong S, Pokaew P, Traisathit P, Srithanaviboonchai K, et al. Tuberculosis treatment with mobile-phone medication reminders in northern Thailand. Southeast Asian J Trop Med Public Health 2011 Nov;42(6):1444-1451. [Medline: 22299414]
6. Story A, Aldridge RW, Smith CM, Garber E, Hall J, Ferenando G, et al. Smartphone-enabled video-observed versus directly observed treatment for tuberculosis: a multicentre, analyst-blinded, randomised, controlled superiority trial. Lancet 2019 Mar 23;393(10177):1216-1224 [FREE Full text] [doi: 10.1016/S0140-6736(18)32993-3] [Medline: 30799062]
7. Wade VA, Karnon J, Eliott JA, Hiller JE. Home videophones improve direct observation in tuberculosis treatment: a mixed methods evaluation. PLoS One 2012;7(11):e50155 [FREE Full text] [doi: 10.1371/journal.pone.0050155] [Medline: 23226243]
8. Shanks L, Masumbuko EW, Ngoy NM, Maneno M, Bartlett S, Thi SS, et al. Treatment of multidrug-resistant tuberculosis in a remote, conflict-affected area of the Democratic Republic of Congo. Int J Tuberc Lung Dis 2012 Aug;16(8):1066-1068. [doi: 10.5588/ijtld.11.0240] [Medline: 22565108]
9. Market study on telemedicine. European Commission. 2018. URL: https://ec.europa.eu/health/sites/health/files/ehealth/docs/2018_provision_marketstudy_telemedicine_en.pdf [accessed 2020-10-07]
10. Kruse C, Betancourt J, Ortiz S, Valdes Luna SM, Bamrah IK, Segovia N. Barriers to the use of mobile health in improving health outcomes in developing countries: systematic review. J Med Internet Res 2019 Oct 09;21(10):e13263 [FREE Full text] [doi: 10.2196/13263] [Medline: 31593543]
11. Scott R, Mars M. Telehealth in the developing world: current status and future prospects. Smart Homecare Technol Telehealth 2015 Feb;3:25-37. [doi: 10.2147/SHTT.S75184]
12. Iribarren SJ, Schnall R, Stone PW, Carballo-Diéguez A. Smartphone applications to support tuberculosis prevention and treatment: review and evaluation. JMIR Mhealth Uhealth 2016 May 13;4(2):e25 [FREE Full text] [doi: 10.2196/mhealth.5022] [Medline: 27177591]
13. McQuarrie EF. The Market Research Toolbox: A Concise Guide for Beginners. 3rd Edition. Thousand Oaks: Sage Publications; 2015.
14. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
15. Pirolli P, Mohan S, Venkatakrishnan A, Nelson L, Silva M, Springer A. Implementation intention and reminder effects on behavior change in a mobile health system: a predictive cognitive model. J Med Internet Res 2017 Nov 30;19(11):e397 [FREE Full text] [doi: 10.2196/jmir.8217] [Medline: 29191800]
16. Lycett HJ, Raebel EM, Wildman EK, Guitart J, Kenny T, Sherlock J, et al. Theory-based digital interventions to improve asthma self-management outcomes: systematic review. J Med Internet Res 2018 Dec 12;20(12):e293 [FREE Full text] [doi: 10.2196/jmir.9666] [Medline: 30541741]
17. Morgan DL. Focus Groups as Qualitative Research. Thousand Oaks: Sage Publications; 1996.
18. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
19. Kalafatis S, Pollard M, East R, Tsogas M. Green marketing and Ajzen's theory of planned behaviour: a cross-market examination. J Consumer Mark 1999;16:460. [doi: 10.1108/07363769910289550]
20. van der Werf MJ, Langendam MW, Huitric E, Manissero D. Knowledge of tuberculosis-treatment prescription of health workers: a systematic review. Eur Respir J 2012 May;39(5):1248-1255. [doi: 10.1183/09031936.00125611] [Medline: 22183482]
21. Khan JA, Irfan M, Zaki A, Beg M, Hussain SF, Rizvi N. Knowledge, attitude and misconceptions regarding tuberculosis in Pakistani patients. J Pak Med Assoc 2006 May;56(5):211-214 [FREE Full text] [Medline: 16767946]
22. Wandwalo ER, Mørkve O. Knowledge of disease and treatment among tuberculosis patients in Mwanza, Tanzania. Int J Tuberc Lung Dis 2000 Nov;4(11):1041-1046 [FREE Full text] [Medline: 11092716]

23. Liam CK, Lim KH, Wong CM, Tang BG. Attitudes and knowledge of newly diagnosed tuberculosis patients regarding the disease, and factors affecting treatment compliance. Int J Tuberc Lung Dis 1999 Apr;3(4):300-309 [FREE Full text] [Medline: 10206500]
24. Keutzer L, Wicha SG, Simonsson US. Mobile health apps for improvement of tuberculosis treatment: descriptive review. JMIR Mhealth Uhealth 2020 Apr 21;8(4):e17246 [FREE Full text] [doi: 10.2196/17246] [Medline: 32314977]
25. Serapelwane MG, Davhana-Maselesele M, Masilo GM. Experiences of patients having tuberculosis (TB) regarding the use of Directly Observed Treatment Short-Course (DOTS) in the North West Province, South Africa. Curationis 2016 Sep 27;39(1):e1-e9 [FREE Full text] [doi: 10.4102/curationis.v39i1.1629] [Medline: 27796102]
26. Ibrahim LM, Hadjia IS, Nguku P, Waziri NE, Akhimien MO, Patrobas P, et al. Health care workers' knowledge and attitude towards TB patients under direct observation of treatment in Plateau state Nigeria, 2011. Pan Afr Med J 2014;18 Suppl 1:8 [FREE Full text] [doi: 10.11694/pamj.supp.2014.18.1.3408] [Medline: 25328627]
27. da Trindade de Andrade E, Hennington EA, Siqueira HRD, Rolla VC, Mannarino C. Perspectives of patients, doctors and medical students at a public university hospital in Rio de Janeiro regarding tuberculosis and therapeutic adherence. PLoS One 2015;10(9):e0137572 [FREE Full text] [doi: 10.1371/journal.pone.0137572] [Medline: 26360291]
28. Kumar AA, De Costa A, Das A, Srinivasa GA, D'Souza G, Rodrigues R. Mobile health for tuberculosis management in south India: is video-based directly observed treatment an acceptable alternative? JMIR Mhealth Uhealth 2019 Apr 03;7(4):e11687 [FREE Full text] [doi: 10.2196/11687] [Medline: 30942696]
29. Zou G, King R, Walley J, Yin J, Sun Q, Wei X. Barriers to hospital and tuberculosis programme collaboration in China: context matters. Glob Health Action 2015;8:27067 [FREE Full text] [doi: 10.3402/gha.v8.27067] [Medline: 26408404]
30. Garfein RS, Collins K, Muñoz F, Moser K, Cerecer-Callu P, Raab F, et al. Feasibility of tuberculosis treatment monitoring by video directly observed therapy: a binational pilot study. Int J Tuberc Lung Dis 2015 Sep;19(9):1057-1064 [FREE Full text] [doi: 10.5588/ijtld.14.0923] [Medline: 26260824]
31. Nhavoto JA, Grönlund A, Klein GO. Mobile health treatment support intervention for HIV and tuberculosis in Mozambique: perspectives of patients and healthcare workers. PLoS One 2017;12(4):e0176051 [FREE Full text] [doi: 10.1371/journal.pone.0176051] [Medline: 28419149]
32. Damhus CS, Emme C, Hansen H. Barriers and enablers of COPD telerehabilitation: a frontline staff perspective. Int J Chron Obstruct Pulmon Dis 2018;13:2473-2482 [FREE Full text] [doi: 10.2147/COPD.S167501] [Medline: 30154650]
33. AlBar AM, Hoque MR. Patient acceptance of e-health services in Saudi Arabia: an integrative perspective. Telemed J E Health 2019 Sep;25(9):847-852. [doi: 10.1089/tmj.2018.0107] [Medline: 30452333]

Abbreviations

COREQ: consolidated criteria for reporting qualitative research
TB: tuberculosis

Edited by G Eysenbach; submitted 07.02.20; peer-reviewed by R Mpofu, S Lin; comments to author 26.03.20; revised version received 20.05.20; accepted 26.07.20; published 02.11.20.

Please cite as:
Margineanu I, Louka C, Vincenti-Gonzalez M, Saktiawati AMI, Schierle J, Abass KM, Akkerman O, Alffenaar JW, Ranchor AV, Stienstra Y
Patients and Medical Staff Attitudes Toward the Future Inclusion of eHealth in Tuberculosis Management: Perspectives From Six Countries Evaluated using a Qualitative Framework
JMIR Mhealth Uhealth 2020;8(11):e18156
URL: https://mhealth.jmir.org/2020/11/e18156
doi:10.2196/18156
PMID:33136052

Original Paper

Mobile Health App for Prostate Cancer Patients on Androgen Deprivation Therapy: Qualitative Usability Study

Junaid Nabi[1,2], MD, MPH; Eugene B Cone[1,2], MD; Anjali Vasavada[1], BSc; Maxine Sun[1], MPH, PhD; Kerry L Kilbridge[3], MD, MSc; Adam S Kibel[1], MD; Donna L Berry[4], PhD, RN; Quoc-Dien Trinh[1,2], MD

[1]Division of Urological Surgery, Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

[2]Center for Surgery and Public Health, Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

[3]Lank Center for Genitourinary Oncology, Dana-Farber Cancer Institute, Boston, MA, United States

[4]Phyllis F. Cantor Center, Dana-Farber Cancer Institute, Boston, MA, United States

Corresponding Author:
Quoc-Dien Trinh, MD
Center for Surgery and Public Health
Department of Surgery
Brigham and Women's Hospital, Harvard Medical School
45 Francis St
ASB II-3
Boston, MA, 02115
United States
Phone: 1 617 525 7350
Email: trinh.qd@gmail.com

Abstract

Background: Androgen deprivation therapy (ADT) increases the risk of metabolic adverse effects among patients with prostate cancer. The transformative impact of mobile health (mHealth) apps may benefit men managing activity and nutrition at home.

Objective: This study aimed to evaluate the usability and patient experience of a newly developed mHealth app among prostate cancer patients on ADT and physicians' beliefs about the potential benefits of using this app.

Methods: This study took place over 2 months, beginning in March 2019. A sample of 5 patients (age 45-75 years) initiating ADT participated in a semistructured focus group discussion with a facilitator. The study participants also included 5 specialist physicians who provided in-depth interviews. An institutional review board–approved script was used to guide both the focus group and physician interviews. Usability was tested through specific scenarios presented to the patients, including downloading the mHealth app, entering information on physical activity and meals, and navigating the app. The focus group and interviews were audio recorded and transcribed. Content analysis was used to analyze the transcripts iteratively and exhaustively. Thematic discrepancies between reviewers were resolved through consensus.

Results: The mean age of the patients was 62 years. This group included 4 White and 1 Latin American patients. The physician specialists included 2 urologists, 2 medical oncologists, and 1 radiation oncologist. Analyses revealed that the patients appreciated the holistic care enabled by the app. Difficulties were observed with registration of the app among 60% (3/5) of the patients; however, all the patients were able to input information about their physical activity and navigate the options within the app. Most patients (4/5, 80%) were able to input data on their recent meal. Among the health care physicians, the dominant themes reflected in the interviews included undermining of patients ability to use technology, patients' fear of technology, and concern for the ability of older patients to access technology.

Conclusions: The patients reported an overall positive experience of using an mHealth app to record and track diet and exercise. Usability was observed to be an important factor for adoption and was determined by ease of registration and use, intuitive appearance of the app, and focus on holistic cancer care. The physicians believed that the app was easy to use but raised concerns about usability among older men who may not typically use smartphone apps.

(JMIR Mhealth Uhealth 2020;8(11):e20224) doi:10.2196/20224

KEYWORDS

mobile health application; prostate cancer; androgen deprivation therapy; thematic analysis; qualitative methods

Introduction

Health care is more efficient and effective when patients are actively engaged in their treatment [1]. Physicians and institutions perceive mobile health (mHealth) technologies—defined as the use of portable devices such as smartphones or tablets for health purposes—as an ideal tool to engage their patients [2]. This technology has already had a transformative impact on the delivery of care [3]. For example, home monitoring has been shown to be effective in reducing cancer symptom distress and improving blood pressure control [4,5]. Additionally, measurements obtained at home can be transmitted wirelessly to electronic medical records or directly to the clinician, allowing for rapid feedback and timely office visits triggered by abnormal values [6].

mHealth technologies are currently being used for a variety of medical conditions, including diabetes, asthma, and chronic obstructive pulmonary disease. With features such as frequent monitoring of patients, active collection of data, and seamless transmission of clinical information into electronic medical records, mHealth apps can enable real-time feedback and lead to improved communication with health care providers [6]. These aspects present a unique opportunity to use mHealth technology as a tool to prevent or alleviate adverse effects of disease and treatment, especially among prostate cancer patients.

While androgen deprivation therapy (ADT) is a standard treatment regimen for intermediate- or high-risk localized disease and metastatic or recurrent prostate cancer, the ensuing hypogonadal state results in significant metabolic and cardiovascular adverse events. Previous studies have documented that men exposed to gonadotropin-releasing hormone agonists—a common form of ADT—manifested a weight gain of 2.4%, an increase in body fat percentage of 9.4%, and a loss in lean body mass of 2%-4% (also known as sarcopenia) by the end of the first year [7,8]. We developed an mHealth app that provides dietary and exercise interventions to monitor and mitigate these metabolic disturbances. The objective of this study was to heuristically evaluate the usability and patient experience of this newly developed mHealth app among prostate cancer patients on ADT and physicians' beliefs about the potential benefits of using this app.

Methods

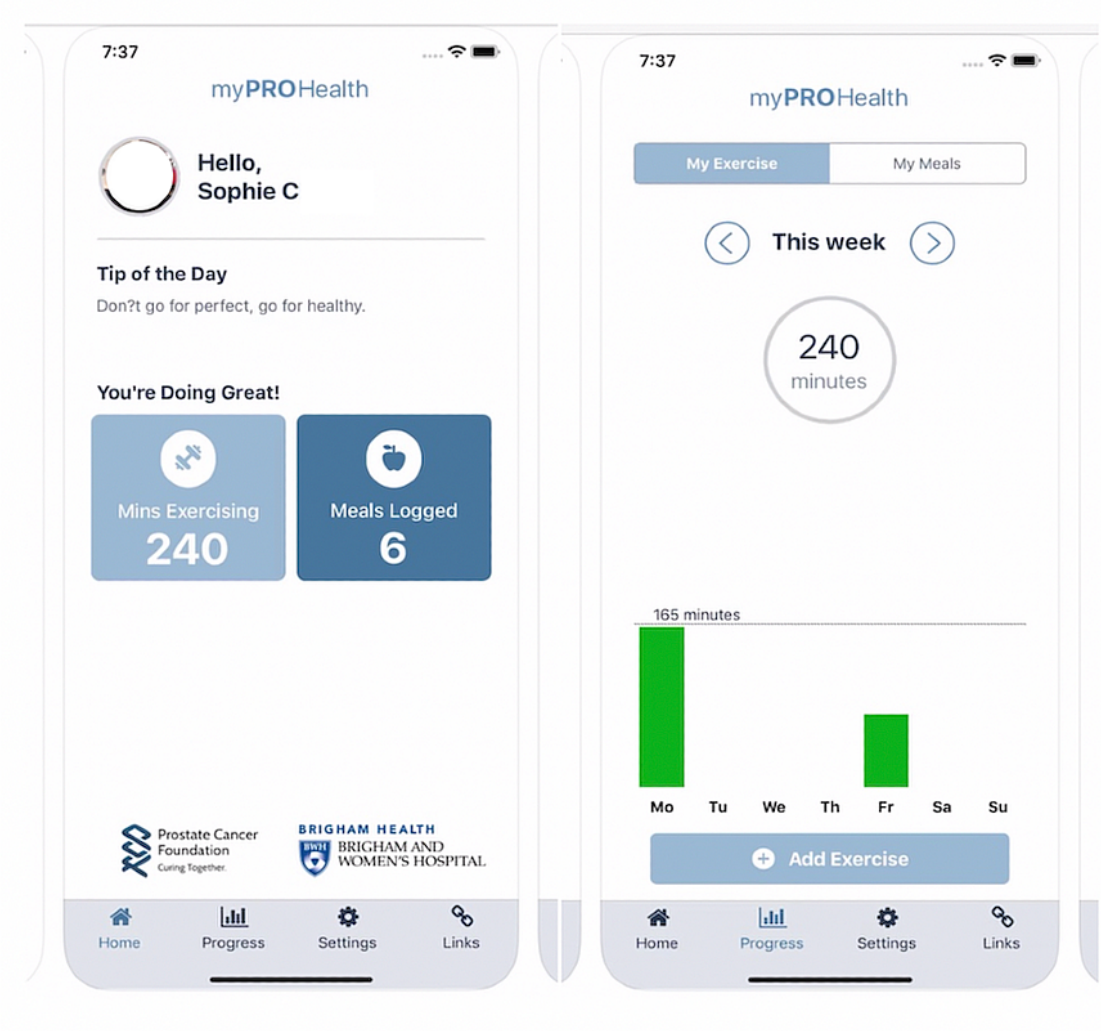
Study Design

This usability study employed qualitative data collection and thematic content analysis methods. Focus group discussions and in-depth interviews were used for collecting data. An institutional review board (IRB)–approved script was used to guide these discussions. Thematic content analysis involves systematically coding and classifying textual information to explicate trends, patterns, frequency, of words, and their embedded relationships and discursive structures [9,10]. This technique enables researchers to understand the attributes of the content by lowering the level of interpretative complexity [11]. The Dana-Farber/Brigham and Women's Cancer Center (DF/BWHCC) IRB approved this study.

The mHealth App

myPROHealth is a health and fitness tracking app developed by the Division of Urological Surgery at Brigham and Women's Hospital. The app seeks to track, record, and promote exercise and good eating habits in men diagnosed with prostate cancer, who are about to initiate ADT. Figure 1 shows sample mHealth app interfaces. The features of the app enable patients to enter information on their recent physical activity and meals as well as monitor their progress.

Figure 1. Screenshots of the mobile health application.



Study Setting and Focus Group Participants

Men diagnosed with prostate cancer, presenting to DF/BWHCC and initiating ADT, were invited to participate. Eligibility criteria included age between 45 and 75 years, the ability to walk 400 m, medical clearance from their primary physician, ability to speak English, cognitive alertness (current Eastern Cooperative Oncology Group performance status available in the medical notes), sufficiently literate (education level available in the medical records), and ownership of a smartphone for >1 year. Men participating in rigorous structured exercise regimes (ie, seeing a professional trainer or training >5 days per week) in the last 6 months were excluded, as their inclusion could conflict with the app's usability for documenting physical activity. Men with planned systemic chemotherapy, planned treatment with abiraterone or enzalutamide, bone metastases, acute illness, any musculoskeletal, cardiovascular, or neurologic disorders that would prevent them from exercising or put them at risk from exercising were excluded, as were men with hearing or visual impairment that would prevent them from participating in the focus group.

Recruitment

Recognizing that the choice of an app-based intervention may skew the sample toward younger and well-educated patients, we provided access to the program for all patients at a cancer clinic as the point of service to reduce potential selection bias. This study took place over 2 months, beginning in March 2019. A convenience sample of 5 patients aged 45-75 years was recruited from the cancer clinic for a focus group discussion. Patients who were initiating ADT were recruited from 3 academic medical clinics affiliated with DF/BWHCC. Permission was sought from the patient's primary care physician or oncologist before contacting them via phone. A research coordinator explained the rationale of the study, and the coordinator obtained verbal consent. The focus group interview was held at a mutually convenient time for all the participants in a hospital conference room. An additional sample of 5 physicians was recruited based on their specialties. Physicians who treat patients undergoing ADT were specifically contacted for individual in-depth interviews. The physicians were selected to gain insights on their beliefs about potential benefits of the mHealth app and whether they would routinely recommend it to their patients if it were readily available.

Patient Focus Group

For the focus group, we used a pilot version of the app, which was available online on Google Play for Android phones and Apple Store for iPhones. Using the feedback from the usability testing, the developers later launched an updated version of the

app. The coordinator followed a pre-established IRB–approved script to guide the focus group. Before initiating the discussion, the group introduced themselves, and then the study coordinator described the goals and expectations of the research project while asking clarifying questions (see Multimedia Appendix 1 for interview script). The focus group discussion was audio recorded and subsequently transcribed with the assistance of a transcription software, and the transcriptions were analyzed manually.

Testing Usability and Definitions of Success

App usability was assessed through the participants' responses while the focus group addressed specific scenarios. These scenarios included participants' ability to register an online account and download the mHealth app, navigate options within the app, appreciate user-interface and engagement with the graphics of the mHealth app (worksheet available in Multimedia Appendix 1), and input information on their (1) typical physical activity in a given week, (2) recent exercise activity, and (3) recent meals (including the ability to take photos of their meals). All the participants used their own smartphones. Success was defined as the ability of the participants to complete the instructions in a given scenario. When instructions were executed completely, the response was considered "successful"; else the response was considered "unsuccessful."

Physician Interviews

Five physicians who treat patients with prostate cancer were identified and contacted. A multi-disciplinary physician cohort well versed in urology, medical oncology, and radiation oncology was recruited. In-depth interviews were scheduled and conducted in locations convenient for both the coordinator and the participant. The same IRB–approved script was used for each physician interview. The interviews were audio recorded and transcribed.

Qualitative Analysis

The developed initial coding scheme was based on usability metrics from previous studies [12,13]. The transcripts were correlated with audio recordings of the patient focus group and physician interviews, and minor edits were made to the transcript (small talk or unrelated dialogue was removed). Once all the data were assembled, a master database of transcripts was compiled and read multiple times by JN and AV. A thematic content analysis was conducted. Significant and recurring themes (defined as themes appearing more than 3 times) were identified and coded in each transcript. Thematic discrepancies were resolved through consensus between the two coders. Each interview and focus group were summarized using these codes. These summaries produced recurring themes across all the data, which allowed us to iteratively refine the codebook and concept map. This cycle continued until all the significant themes were categorized and the transcripts and codebook were consistent with those themes.

Results

Baseline Characteristics

Table 1 shows the baseline characteristics of the patients and physicians recruited for this study.

Table 1. Participants' characteristics (N=5).

Characteristics	Value
Patients (N=5)	
Age (years), mean	62
Gender – male, n (%)	5 (100)
Age (years), n (%)	
45-55	1 (20)
56-65	2 (40)
66-75	2 (40)
Medical conditions, n (%)	
Prostate cancer	5 (100)
Race/ethnicity, n (%)	
White	4 (80)
Hispanic/Latino	1 (20)
Education status, n (%)	
Graduate/professional degree	2 (40)
Some college	2 (40)
GED[a]/high school diploma	1 (20)
Physicians (N=5)	
Specialty, n (%)	
Urologist	2 (40)
Medical oncologist	2 (40)
Radiation oncologist	1 (20)

[a]GED: General education development.

Patient Focus Group Results

The focus group discussion with the participants covered the following dominant patient themes surrounding mHealth app usability: limited knowledge, participants' appreciation for holistic care, and struggles with app registration and download. The focus group discussion also revealed that participants' familiarity with mHealth apps was limited. This aspect was highlighted by statements such as, "There are so many options, wow." Participants highly appreciated that health care organizations and physicians were proactively building digital solutions that would improve patient well-being and convenience, as evidenced by statements such as, "This sort of lets the patient know that there is more to us than just the cancer, there's also our general well-being." Last, participants were candid about the struggles they experienced while handling mHealth apps in general as patients with cancer, as revealed by statements such as, "It might just be me but it's tough for me to understand." Table 2 presents detailed statements on these themes.

Table 2. Participants' reflections on the usage of the mHealth app.

Theme	Quotations
Participant knowledge	• "There are so many options, wow – I didn't realize that things that I do on a daily basis would be considered exercise. My wife would be so happy to hear that! What else did you ask? If it was intuitive? Yes, it is clear and intuitive to me."
Participant appreciation	• "I think it's great that Dr.———, and whoever else is working on this study, cares about the—you know holistic care of the patient. This sort of lets the patient know that there is more to us than just the cancer, there's also our general well-being." • "No, this is great. It shows that the providers care and that we are more than a number to them, I think."
Participant struggling	• "To be honest, I don't think this is intuitive. It might just be me but it's tough for me to understand."

Participants' responses to specific scenarios revealed that they were generally comfortable using the mHealth app for documenting exercise and diet. For Scenario 1, which tested the participants' ability to register an online account and download the mHealth app, we found that 60% (3/5) were "successful" and 40% (2/5) were "unsuccessful." For Scenario 2, which tested the participants' ability to input information on their typical physical activity in a given week, we found that 100% (5/5) were "successful." For Scenario 3, which tested the participants' ability to navigate options within the app, we found that 100% (5/5) were "successful." For Scenario 4, which tested the participants' ability to input information on their recent exercise activity, we found that 60% (3/5) were "successful" and 40% (2/5) were "unsuccessful." For Scenario 5, which tested the participants' ability to input information on their recent meals as well as take photos of their meals, we found that 80% (4/5) were "successful" and 20% (1/5) were "unsuccessful." Last, for Scenario 6, which tested the participants' ability to appreciate the user-interface and engage with the graphics of the mHealth app, we found that 100% (5/5) were "successful."

Physician Interview Results

Among the health care physicians, the dominant themes reflected their concern regarding the undermining of patients, patients' fear of technology, and the ability of older patients to access a smartphone app (Table 3). When queried about patients' ability to employ technology for recording exercise and dietary activities, the physicians believed that most patients would not be able to do so, as expressed by their statements:

> *I mean, I bet some patients are tech savvy, but the majority of patients I see are not.*

The interviews also revealed that the physicians were apprehensive that integrating technology as part of their health care would lead to increased stress among their patients:

> *I think patients will struggle with this. I think that we need to be careful with technological…interventions like this because it may make them more stressed.*

We also found that the physicians believed that age would prevent older patients from benefitting from the advantages of using the mHealth app:

> *But I worry that the patients won't actually use it. It might be hard for older patients, and most prostate cancer patients are above the age of 45.*

In addition to these impressions, the physicians also emphasized the intuitiveness and ease of navigating the app, as indicated by statements such as, "…it's easy to go back and forth between tabs" and "It's a great app. The design is simple, intuitive, and professional."

Table 3. Physicians' reflections on the usage of the mHealth app.

Theme	Quotations
Undermining patients	• "I think my con is that, I don't think older patients, or any patients with prostate cancer will understand how to use this app." • "I mean, I bet some patients are tech savvy, but the majority of patients I see are not."
Patients' fear of technology	• "I think patients will struggle with this. I think that we need to be careful with technological…interventions like this because it may make them more stressed."
Concerns for older patients	• "The app is really well designed and is…generally pretty intuitive. But I worry that the patients won't actually use it. It might be hard for older patients, and most prostate cancer patients are above the age of 45."

Discussion

Findings

Mobile health apps present a unique opportunity to enhance patient engagement and self-management of chronic diseases. In this study, we found that the overall usability of an mHealth app in patients with prostate cancer participants on ADT was acceptable. Moreover, we found that participants generally appreciated their care teams' efforts in devising technological solutions that would enhance their ability to record and monitor their diet and exercise. We also observed that physicians were generally skeptical of the benefits of integrating technology with conventional health care to augment prostate cancer care. A disconnect was found between the experiences of the participants and the physicians on the potential of health technology solutions for improving prostate cancer care.

The findings of our study have several implications. First, we found that the participants were generally successful in using the mHealth app and were able to input data about their meals and exercise. Most participants were able to navigate through various features of the app (ie, move between the tabs and the home screen) and take photos of their meals. Our findings contrast those of Sarkar et al [14], who found that the usability and applicability of mHealth apps for self-management of chronic conditions was limited [14]. Sarkar et al [14] examined the acceptability and usability of mHealth apps in a set of racially diverse and economically disadvantaged patients—communities that often have higher prevalence of chronic diseases (diabetes and depression) and require caregiving. They also examined usability by asking participants to perform specific tasks on 11 apps. While our study also examined usability by asking participants to perform a variety of tasks, we had a more specific focus, namely, the participants' ability to record exercise and diet, interventions that can impact treatment, and disease outcomes. In addition, Sarkar et al [14] attempted to understand usability among existing apps, while our goal was to document how a newly developed app can be improved for greater acceptance among patients. We found that the participants appreciated that they could use the mHealth app for tracking their typical physical activity—an important consideration for patients who may develop sarcopenia. This

finding correlates with those of previous studies reporting that prostate cancer patients are often receptive to using mHealth apps to promote physical activity, especially if the technology is easy to use [15]. ADT is associated with decreased bone mineral density and higher risk of fracture, and current evidence indicates that improvements in dietary intake (calcium and vitamin D supplementation) could alleviate these risks [16]. Our analysis underscores the benefit of using mHealth apps to record and track dietary intake. Additionally, patients on ADT are at risk of developing several metabolic adverse events, including weight gain and diabetes [7,16]. Moreover, given that mHealth apps can record the level of physical activity, these online tools can potentially mitigate the impact of metabolic adverse events.

Second, we found that our sample of physicians was not convinced that using mHealth technology for prostate cancer patients would be beneficial, although they remarked positively on its general usability, appearance, and navigation. These perceptions were based on their beliefs that patients lacked technological understanding, would be unable to use mHealth apps, and their advanced age would preclude them from benefitting from technological interventions. This finding is in conflict with those of a previous work as well as the results of patients' usability evaluations from this study. Nguyen et al [17] reported that physicians viewed patient-focused apps positively and believed that these interventions would augment their ability to provide patient care. The same study also observed that physicians believed that using mHealth apps would supplement their role by providing additional medical information to their patients electronically and therefore enhance self-management of chronic conditions.

Third, our findings point to the role of physician specialization as a determinant of mHealth app adoption. As per Bodur et al [18], medical and nursing students believed that health technology could play an increasingly important role in the delivery of care in the future [18]. It is possible that these differences are driven by variations in clinical focus; we interviewed prostate cancer specialists, and previous studies have mostly evaluated general practitioners. Although there is evidence that cancer specialists are highly satisfied with using mHealth technology for improving cancer care delivery, it seems that specialists perceive a greater benefit of adopting mHealth apps at the population, rather than the patient, level [19].

Last, our analyses also reveal the various factors that can influence the adoption of mHealth apps. Our qualitative analyses underscore how ease of use and registration/download, intuitive appearance of the app, ability to navigate various tabs within the app, and focus on holistic cancer care are important adoption considerations for patients as well as physicians. These findings concur with previous analyses. In a comprehensive review, Alam et al [20] reported that factors such as performance expectancy (the belief that technology will achieve expected outcomes), effort expectancy (ease of use), and facilitating conditions (the belief that technology/organization infrastructure supports usage) were positively associated with greater adoption of mHealth technology [20]. Our study provides additional evidence to support incorporation of these usability features in mHealth apps for patients.

Given the methodological approach we pursued, our study has several limitations. We studied a small sample of physicians and patients from a single cancer center. The results may vary with larger–scale implementation. We evaluated the participants in our study on a specific mHealth app and provided the service free of charge, which may lead to selection bias and limit the generalizability of this work, respectively. While some mHealth apps provide information on prostate cancer, a recent systematic review revealed that most commercially available apps do not provide culturally sensitive information and have major usability concerns [21]. Other studies have reported that patients would be willing to pay for these services—especially if they are affordable [22,23]. In our study, the patient participants were aware that their responses were being recorded and evaluated, which could have led to changes in observed behavior [24] and overestimated the usability of the app due to social desirability bias. Although the participants were aware that their own care teams were involved in this study, results may vary when implementing such mHealth apps in a different context. Moreover, we did not evaluate the efficacy of the app to improve prostate cancer outcomes; we chose to first evaluate the usability and feasibility of the app, as we viewed these aspects as necessary requirements for future studies on efficacy evaluation.

Conclusion

Usability was observed to be an important factor for adoption of the proposed mHealth technology, as determined by ease of registration and use, intuitive appearance of the app, and focus on holistic cancer care. Men receiving ADT for prostate cancer treatment had an overall positive experience using the mHealth app to record and track their diet and exercise. The physicians believed that the app was easy to use but raised concerns about usability among older men who do not commonly use smartphone apps. This work provides foundational evidence for studies that examine the feasibility, efficacy, and interoperability of using mHealth apps for prostate cancer patients.

Acknowledgments

This project was funded by the 2016 Clay Hamlin Young Investigator Award (16YOUN20), Prostate Cancer Foundation. QDT received research funding from the Health Services Research pilot test grant, Defense Health Agency (HU0001-18-2-0018), and an unrestricted educational grant from the Vattikuti Urology Institute.

Conflicts of Interest

QDT reports personal fees from Astellas, Bayer, Janssen, and Insightec.

Multimedia Appendix 1
Usability test.
[PDF File (Adobe PDF File), 18 KB - mhealth_v8i11e20224_app1.pdf]

References

1. Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. Health Aff (Millwood) 2015 Mar;34(3):431-437. [doi: 10.1377/hlthaff.2014.0452] [Medline: 25732493]
2. Broderick A, Haque F. Mobile health and patient engagement in the safety net: a survey of community health centers and clinics. Issue Brief (Commonw Fund) 2015 May;9:1-9. [Medline: 26040018]
3. Cortez NG, Cohen IG, Kesselheim AS. FDA regulation of mobile health technologies. N Engl J Med 2014 Jul 24;371(4):372-379. [doi: 10.1056/NEJMhle1403384] [Medline: 25054722]
4. Berry DL, Hong F, Halpenny B, Partridge AH, Fann JR, Wolpin S, et al. Electronic self-report assessment for cancer and self-care support: results of a multicenter randomized trial. J Clin Oncol 2014 Jan 20;32(3):199-205 [FREE Full text] [doi: 10.1200/JCO.2013.48.6662] [Medline: 24344222]
5. Plante TB, Urrea B, MacFarlane ZT, Blumenthal RS, Miller ER, Appel LJ, et al. Validation of the instant blood pressure smartphone app. JAMA Intern Med 2016 May 01;176(5):700-702 [FREE Full text] [doi: 10.1001/jamainternmed.2016.0157] [Medline: 26938174]
6. Steinhubl SR, Muse ED, Topol EJ. Can mobile health technologies transform health care? JAMA 2013 Dec 11;310(22):2395-2396. [doi: 10.1001/jama.2013.281078] [Medline: 24158428]
7. Smith MR, Finkelstein JS, McGovern FJ, Zietman AL, Fallon MA, Schoenfeld DA, et al. Changes in body composition during androgen deprivation therapy for prostate cancer. J Clin Endocrinol Metab 2002 Feb;87(2):599-603. [doi: 10.1210/jcem.87.2.8299] [Medline: 11836291]
8. Dawson JK, Dorff TB, Todd Schroeder E, Lane CJ, Gross ME, Dieli-Conwright CM. Impact of resistance training on body composition and metabolic syndrome variables during androgen deprivation therapy for prostate cancer: a pilot randomized controlled trial. BMC Cancer 2018 Apr 03;18(1):368 [FREE Full text] [doi: 10.1186/s12885-018-4306-9] [Medline: 29614993]
9. Grbich C. Qualitative data analysis: An introduction. In: Researching Practice. Boston (USA Office): Brill; Jan 10, 2010:173-183.
10. Morgan DL. Qualitative content analysis: a guide to paths not taken. Qual Health Res 1993 Feb;3(1):112-121. [doi: 10.1177/104973239300300107] [Medline: 8457790]
11. Bloor M, Wood F. Keywords in qualitative methods: A vocabulary of research concepts. London, UK: SAGE Publications; 2006:A.
12. Taha J, Sharit J, Czaja SJ. The impact of numeracy ability and technology skills on older adults' performance of health management tasks using a patient portal. J Appl Gerontol 2014 Jun;33(4):416-436 [FREE Full text] [doi: 10.1177/0733464812447283] [Medline: 24781964]
13. Segall N, Saville JG, L'Engle P, Carlson B, Wright MC, Schulman K, et al. Usability evaluation of a personal health record. AMIA Annu Symp Proc 2011;2011:1233-1242 [FREE Full text] [Medline: 22195184]
14. Sarkar U, Gourley GI, Lyles CR, Tieu L, Clarity C, Newmark L, et al. Usability of commercially available mobile applications for diverse patients. J Gen Intern Med 2016 Dec;31(12):1417-1426 [FREE Full text] [doi: 10.1007/s11606-016-3771-6] [Medline: 27418347]
15. Roberts AL, Potts HW, Koutoukidis DA, Smith L, Fisher A. Breast, prostate, and colorectal cancer survivors' experiences of using publicly available physical activity mobile apps: Qualitative study. JMIR Mhealth Uhealth 2019 Jan 04;7(1):e10918 [FREE Full text] [doi: 10.2196/10918] [Medline: 30609982]
16. Nguyen PL, Alibhai SMH, Basaria S, D'Amico AV, Kantoff PW, Keating NL, et al. Adverse effects of androgen deprivation therapy and strategies to mitigate them. Eur Urol 2015 May;67(5):825-836. [doi: 10.1016/j.eururo.2014.07.010] [Medline: 25097095]
17. Nguyen AD, Frensham LJ, Baysari MT, Carland JE, Day RO. Patients' use of mobile health applications: what general practitioners think. Fam Pract 2019 Mar 20;36(2):214-218. [doi: 10.1093/fampra/cmy052] [Medline: 29873708]
18. Bodur G, Gumus S, Gursoy NG. Perceptions of Turkish health professional students toward the effects of the internet of things (IOT) technology in the future. Nurse Educ Today 2019 Aug;79:98-104. [doi: 10.1016/j.nedt.2019.05.007] [Medline: 31112847]
19. Sharma JJ, Gross G, Sharma P. Extending oncology clinical services to rural areas of Texas via teleoncology. J Oncol Pract 2012 Mar;8(2):68 [FREE Full text] [doi: 10.1200/JOP.2011.000436] [Medline: 29447100]
20. Alam MZ, Hoque MR, Hu W, Barua Z. Factors influencing the adoption of mHealth services in a developing country: A patient-centric study. International Journal of Information Management 2020 Feb;50:128-143. [doi: 10.1016/j.ijinfomgt.2019.04.016]
21. Owens OL, Beer JM, Reyes LI, Thomas TL. Systematic review of commercially available mobile phone applications for prostate cancer education. Am J Mens Health 2019;13(1):1557988318816912. [doi: 10.1177/1557988318816912] [Medline: 30526243]

22. Ebner C, Wurm EM, Binder B, Kittler H, Lozzi GP, Massone C, et al. Mobile teledermatology: a feasibility study of 58 subjects using mobile phones. J Telemed Telecare 2008;14(1):2-7. [doi: 10.1258/jtt.2007.070302] [Medline: 18318921]
23. Pathipati AS, Ko JM. Implementation and evaluation of Stanford Health Care direct-care teledermatology program. SAGE Open Med 2016;4:2050312116659089. [doi: 10.1177/2050312116659089] [Medline: 27493756]
24. Borycki E, Monkman H, Griffith J, Kushniruk AW. Mobile usability testing in healthcare: Methodological approaches. Stud Health Technol Inform 2015;216:338-342. [Medline: 26262067]

Abbreviations

ADT: Androgen Deprivation Therapy
DF/BWHCC: Dana-Farber/Brigham and Women's Cancer Center
GED: General Education Development
IRB: Institutional Review Board
mHealth: mobile health

Edited by G Eysenbach; submitted 13.05.20; peer-reviewed by B Chaudhry, C Jacob; comments to author 11.07.20; revised version received 22.08.20; accepted 13.09.20; published 03.11.20.

Please cite as:
Nabi J, Cone EB, Vasavada A, Sun M, Kilbridge KL, Kibel AS, Berry DL, Trinh QD
Mobile Health App for Prostate Cancer Patients on Androgen Deprivation Therapy: Qualitative Usability Study
JMIR Mhealth Uhealth 2020;8(11):e20224
URL: https://mhealth.jmir.org/2020/11/e20224
doi:10.2196/20224
PMID:33141104

Original Paper

A Smartphone-Based Technique to Detect Dynamic User Preferences for Tailoring Behavioral Interventions: Observational Utility Study of Ecological Daily Needs Assessment

Ginger E Nicol[1], MD; Amanda R Ricchio[1], BA; Christopher L Metts[2], MD; Michael D Yingling[1], MS; Alex T Ramsey[3], PhD; Julia A Schweiger[1], CCRC; J Philip Miller[4], BA; Eric J Lenze[1], MD

[1]Healthy Mind Lab, Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, United States

[2]Department of Pathology and Laboratory Medicine, College of Medicine, Medical University of South Carolina, Charleston, SC, United States

[3]Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, United States

[4]Division of Biostatistics, Washington University School of Medicine, St. Louis, MO, United States

Corresponding Author:
Ginger E Nicol, MD
Healthy Mind Lab
Department of Psychiatry
Washington University School of Medicine
600 South Taylor Avenue
Suite 121
St. Louis, MO, 63110
United States
Phone: 1 314 362 5939
Fax: 1 314 747 1160
Email: nicolg@wustl.edu

Abstract

Background: Mobile health apps are promising vehicles for delivering scalable health behavior change interventions to populations that are otherwise difficult to reach and engage, such as young adults with psychiatric conditions. To improve uptake and sustain consumer engagement, mobile health interventions need to be responsive to individuals' needs and preferences, which may change over time. We previously created an ecological daily needs assessment to capture microprocesses influencing user needs and preferences for mobile health treatment adaptation.

Objective: The objective of our study was to test the utility of a needs assessment anchored within a mobile app to capture individualized, contextually relevant user needs and preferences within the framework of a weight management mobile health app.

Methods: Participants with an iOS device could download the study app via the study website or links from social media. In this fully remote study, we screened, obtained informed consent from, and enrolled participants through the mobile app. The mobile health framework included daily health goal setting and self-monitoring, with up to 6 daily prompts to determine in-the-moment needs and preferences for mobile health–assisted health behavior change.

Results: A total of 24 participants downloaded the app and provided e-consent (22 female; 2 male), with 23 participants responding to at least one prompt over 2 weeks. The mean length of engagement was 5.6 (SD 4.7) days, with a mean of 2.8 (1.1) responses per day. We observed individually dynamic needs and preferences, illustrating daily variability within and between individuals. Qualitative feedback indicated preferences for self-adapting features, simplified self-monitoring, and the ability to personalize app-generated message timing and content.

Conclusions: The technique provided an individually dynamic and contextually relevant alternative and complement to traditional needs assessment for assessing individually dynamic user needs and preferences during treatment development or adaptation. The results of this utility study suggest the importance of personalization and learning algorithms for sustaining app engagement in young adults with psychiatric conditions. Further study in broader user populations is needed.

(JMIR Mhealth Uhealth 2020;8(11):e18609) doi:10.2196/18609

KEYWORDS

mobile health; telemedicine; needs assessment; healthy lifestyle; ecological momentary assessment; mobile applications; behavior intervention; behavior therapy

Introduction

Background

Digital health interventions are being developed at a rapid rate, particularly for behavioral management of chronic diseases like diabetes, heart disease, and stroke, which account for greater than 90% of health care expenditures and more than 75% of deaths in the United States [1,2]. Such interventions promise increased access and scalability across health care systems. Mobile health (mHealth) interventions in particular offer immediate engagement, often titratable to user preferences, with the ability to deploy assessments that incorporate both self-report and sensor data. However, approximately half of users disable health apps within 2 weeks of download, with loss of interest a commonly cited reason for app disablement [3]. Barriers to and facilitators of engagement, adherence, and motivation all likely vary over time, impacted by unique, context-dependent needs and preferences both intrinsic and extrinsic to the individual—variables that cannot be adequately assessed by traditional needs assessment methods. Person-centered strategies to optimize patient engagement are critically needed for mHealth apps to impart intended benefits, and for mHealth interventions to become a mainstream way of delivering health care [4].

Obesity treatment is a case where eliciting patient preferences in treatment development is particularly important, as (1) patient engagement is a critical indicator of treatment success, (2) there is a range of treatment options, and (3) there are limited data to guide personalization of a treatment approach [5]. Only 20% to 30% of individuals seeking first-line behavioral weight loss treatment achieve clinically meaningful or sustained weight loss [6,7]. Consistent predictors of treatment response across weight loss studies include treatment engagement (eg, session attendance, homework completion) [8], adherence to self-monitoring activities (eg, food and activity logging, tracking weight) [9], and intrinsic motivation for health behavior change [10]. However, individuals with psychiatric conditions may experience unique and persistent barriers to engagement in healthy eating and exercise [11,12]. For example, parents of adolescents with psychiatric illness who participated in a family-based weight loss intervention reported that major barriers to participation were pragmatic (transportation to and from in-person visits), or were directly related to the psychiatric condition (shame and avoidance). These individuals expressed preferences for digital self-monitoring and support delivered via a mobile device over in-person treatment. Moreover, participants indicated that needs and barriers varied from day to day, and that an intervention adaptive to their dynamically changing needs was important [13]. As the availability of mobile devices among mentally ill youth increases [14], this mode of treatment becomes a feasible way to increase access and engagement [15]. However, to our knowledge, no mHealth approaches to health promotion, including weight management, have been adapted to engage patients with early-onset psychiatric conditions, where prevention is a primary public health concern [16,17].

Traditional behavioral treatment adaptation approaches use a range of methods (eg, semistructured interviews, checklists, self-report questionnaires, focus groups) to elicit needs and preferences from potential recipients to inform treatment development and adaptation [18-20]. However, these methods are often employed at static assessment times, missing the dynamic effects of time and context on user needs and preferences, which are critical for treatment engagement. Digital data collection methods, such as ecological momentary assessment [21] and passively collected sensor data, can inform adaptation efforts. For example, just-in-time adaptive interventions track the dynamic effects of time and context on individual characteristics or behaviors and adapt, ideally in real time, based on a concrete measurable construct such as self-reported mood [22] or sedentary behavior [23]. With enough measurement time points, such ideographic assessment methods can be used to model and predict dynamic changes in symptom presentation or treatment response in a single individual, without the need for reference to a larger group [24,25]. However, few just-in-time adaptive interventions have collected or incorporated dynamic user preferences and perspectives into adaptation algorithms [26,27], and little is known about what types of data inputs are most relevant in building adaptive treatments for lifestyle change [28].

Objective

To begin the process of adapting an existing interactive obesity treatment approach [29,30] for use in young people with psychiatric conditions, we created an ecological daily needs assessment. The intention was to capture dynamic user needs and preferences data as part of the adaptation process for developing a self-adapting treatment algorithm. However, the amount of lived-experience feedback needed to inform relevant treatment adaptation and, more importantly, the threshold for assessment fatigue, is not well understood [31]. The primary aim of this study was to evaluate the utility of the tool over a period of 2 weeks as part of an overall mHealth treatment adaptation effort, specifically to determine whether it could detect dynamic user needs and preferences. Secondarily, we aimed to determine the threshold for response fatigue in young adults with psychiatric conditions.

Methods

Participant Recruitment and Study Orientation

In this fully remote observational utility study, we recruited participants nationally via social media (Twitter, Facebook, and Instagram), as well as traditional methods, using an online research registry (ResearchMatch, Vanderbilt University); US national email listservs for college students, medical students, and residents; and word-of-mouth or flyers. These recruitment methods directed potential participants to the study website in

order to download the free app. Once participants had downloaded the app, they were asked to complete a brief screening questionnaire, which included questions about comfort with and ability to use their device for answering preferences questions. Inclusion criteria were age between 18 and 45 years; access to a web-enabled device with the iPhone operating system (iOS; Apple Inc); ability to keep their device with them for most of the day over the following 2 weeks for the purpose of answering needs assessment prompts; and a history of a diagnosis of a psychiatric or psychological disorder. Since the app involved setting health goals for weight loss, a history of an eating disorder was exclusionary.

After reading about the study and passing the eligibility screen, participants signed consent on their device touch screen and received an email with the full signed consent document, which included email addresses and phone numbers for the study coordinator and principal investigator. Participants were asked about their preferred times for the following questions: (1) a time every morning to set a health intention for the day, (2) times the user typically ate breakfast, lunch, and dinner, or would be most likely to engage in physical activity, and (3) a time in the evening to answer questions about app usability and acceptability. Participants were then prompted 6 times daily each for eating and exercise needs assessments during their preferred times.

Study Approval

This study was reviewed and approved by the Washington University Institutional Review Board and the Washington University Office of Information Security in December 2016. Following completion of development in August 2017, study enrollment was open from September 2017 to March 2018.

Mobile App Development

As the first step in a larger, ongoing research program to adapt and test the effectiveness of an existing weight management intervention [29,30] for use in young adults with severe mental illness, we created a needs assessment tool for the purpose of conducting a digital needs assessment within the mHealth context. The theoretical behavior change framework underlying the overall program is based on increasing self-efficacy [32] for mental and physical health, by using mobile technology to reduce extraneous cognitive load associated with making health behavior changes [33] and to increase uptake of information and intervention engagement. The ecological needs assessment was embedded within a basic mHealth goal-setting framework.

Needs Assessment Description

We developed the needs assessment using Status/Post, a digital platform (developed by CLM) that integrates the Apple ResearchKit (Apple Inc) framework with the REDcap (REDCap Consortium) web app through an application programming interface [34]. The app framework included a daily prompt for health intention setting and scaling, 6 prompts for needs assessment during the day designated by users as times they would be most likely to eat (3 prompts) or exercise (3 prompts), and reflection on goal progress at the end of the day (screen shots in Multimedia Appendix 1). The final survey each day consisted of feasibility, acceptability, and appropriateness questions adapted from Lyon and colleague's contextualized technology adaptation process [35]. Feasibility questions focused on eliciting feedback from the user about the frequency and timing of question prompts, as well as messaging content. We conducted semistructured interviews with participants who had high (eg, responses on >80% of study days) and low (eg, responses on <25% of study days) engagement following completion of the 2-week utility study.

Statistical Analysis

The recommended number of users needed for maximal detection of usability problems is 5 [36], and up to 25 users for comparative studies [37]. Anticipating a 50% attrition rate [3] over the 2-week study period, we enrolled 35 participants with the goal of at least 10 completers. We generated descriptive statistics (mean, frequencies, and proportions) for survey responses. We converted Likert-scale items from severity to a numeric rating (eg, 1 = not at all; 3 = very much). Over the course of the study, there were 235 potential eating prompt responses, comprised of 9 possible response options (ie, "Are you planning to eat? yes/no;" "Are you eating out? yes/no;" needs assessment question with 6 response options; and a free-text option for additional feedback), and 234 possible exercise prompt responses, comprising 8 possible response options ("Are you planning to exercise? yes/no;" needs assessment question with 6 response options; and free-text option for additional feedback). All participants with at least one postbaseline response were included in the analysis. We report responses to each survey question as number (n) and percentage of participants with a response in each category for a given question. We cleaned and analyzed data using the R software package version 3.1.1 2014-07-11, R.app 1.65 (R Foundation for Statistical Computing). We used all available data from all participants.

Results

Participant Characteristics

A total of 35 individuals downloaded the app, 24 consented to participate, and 23 proceeded past the orientation questions to answer at least one needs assessment question; 2 of them were male. Participants accessed the app via the study website (n=4), the online study registry (n=6), social media (n=6), an email listserv (n=6), or a flyer or by word-of-mouth (n=9). Participants were queried 6 times daily: 3 times daily during prespecified 1-hour periods when they were most likely to eat meals, and 3 times daily when they were most likely to exercise. Of the 23 participants who provided responses, 18 responded "Yes" to at least one eating query during the study, prompting the healthy eating needs assessment questions to be deployed 122 times over the study period. A total of 11 participants responded "Yes" to at least one exercise query during the study, prompting healthy activity needs assessment questions to be deployed 28 times during the study period.

App Use and Engagement

The mean length of participation in the study was 5.6 (SD 4.7) days, with mean of 2.8 (1.1) responses per day. The mean number of responses per participant was 6.7 (3.0) over the

2-week study period. The earliest time to termination was 1 day; 2 individuals completed the entire 2 weeks. As Figure 1 shows, app use, defined as the number of participants responding to prompts, decreased by 46% (18/39) during the first day, with a varied but continual decrease in the number of respondents over the course of the study. However, the pooled daily response rate decreased to 80% (31/39) on day 1 and then remained stable at 67% (8/12) on day 7 through day 14 (6/9). Response rates remained relatively stable during the second week of the study, although individual responses varied across study day and assessment time point (Figure 2).

Figure 1. Overall study engagement shown by the number of total responses per day (green line), and "yes" responses per day for eating (orange bars) and exercise (purple bars). The table presents raw numbers of responses per day based on mealtime.

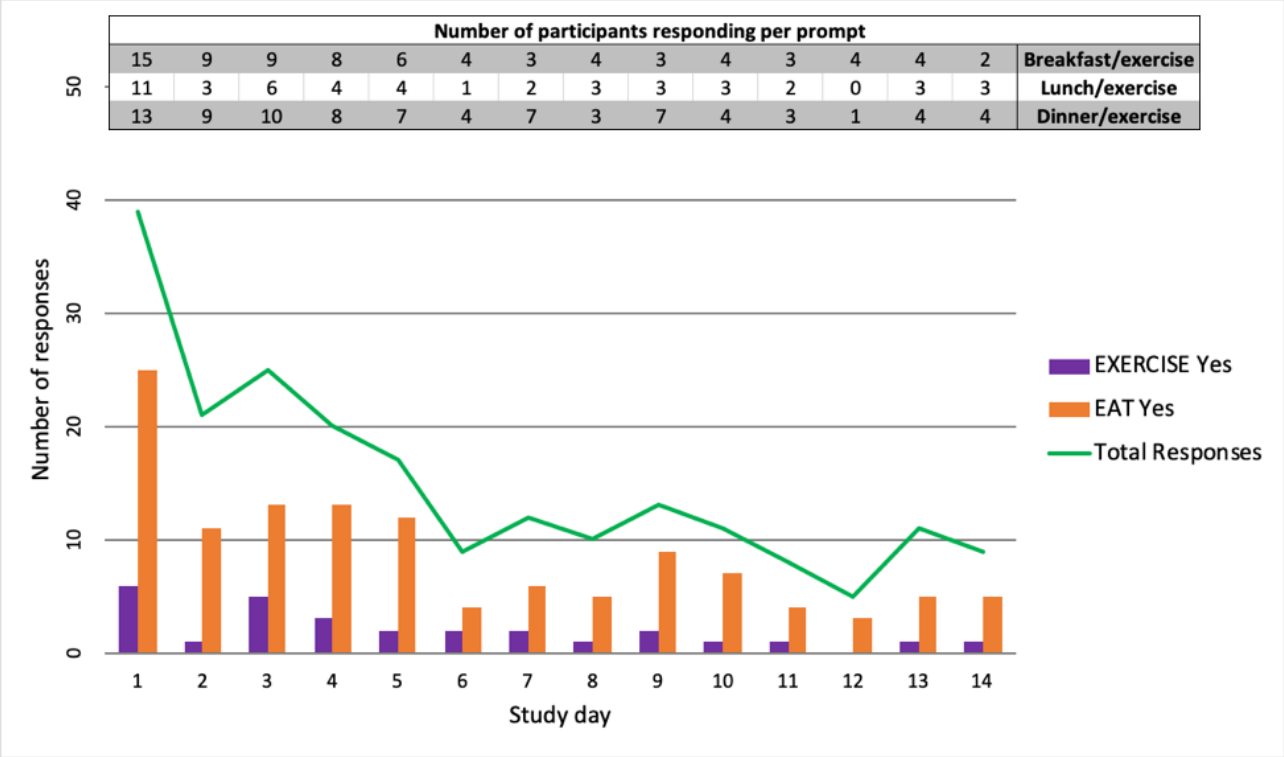
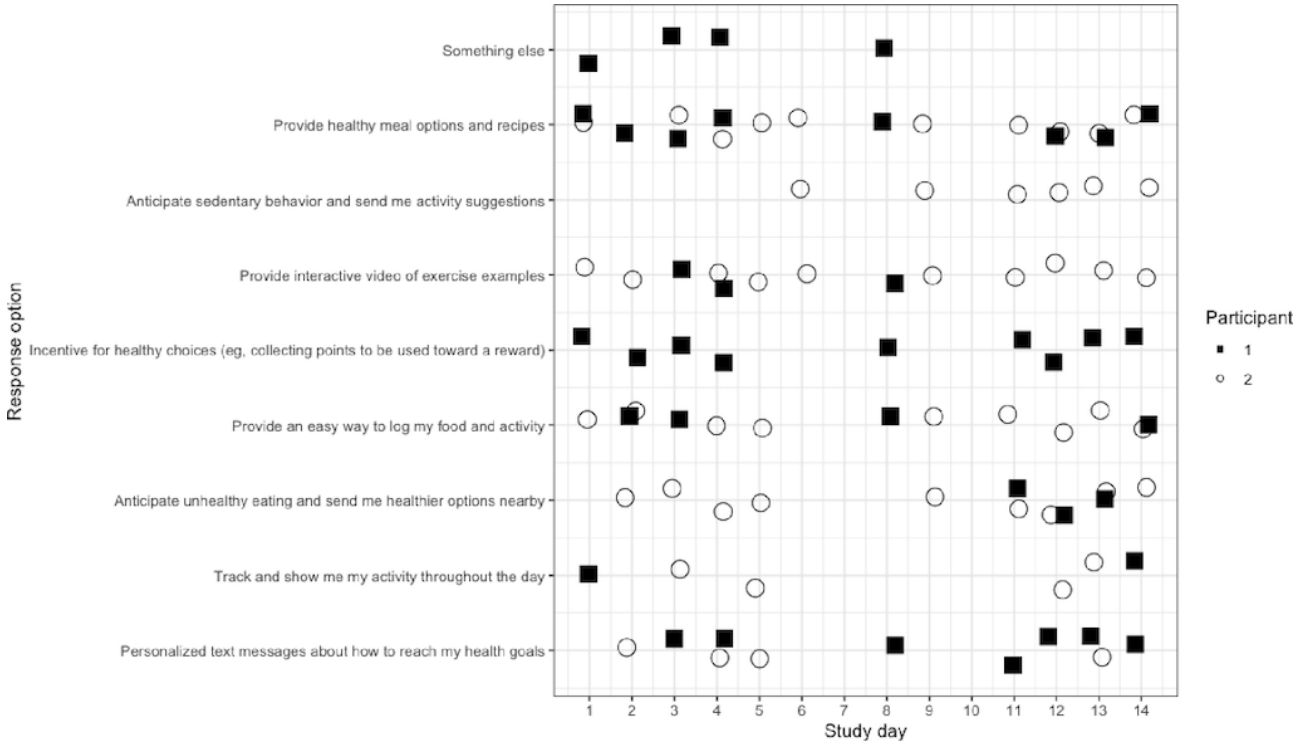


Figure 2. Responses on dynamic needs and preferences for healthy eating from 2 study participants over the 2-week study period.



User Needs and Preferences

As Table 1 shows, the top preferences for app functionality were receiving a reminder of the daily health goal, simplified self-monitoring for food intake, healthy eating and activity options nearby, and being able to see or track progress on goals throughout the day. More nuanced qualitative responses (eg, participants chose "something else" and entered a text response) suggested that participants wanted pragmatic assistance with making healthy eating selections ("Access to healthy preferred options or premade healthy options;" "An inventory of food in my house so I can figure out healthy options;" "Let me list at the beginning of the week what I have on hand and use those as suggestions"). Participants also indicated their preference for the app to learn their patterns and behaviors to anticipate their needs, as well as respond with personalized support ("Motivational messaging and a dashboard of my progress;" "Suggestions on healthy options in my area;" "Give me a list of 3-5 quick and healthy meal options;" "Someone to go with me and hold me accountable;" "Support me by checking in with me like a friend would").

Responses to preference questions also revealed dynamic responses within and between individuals throughout each day over the course of the study (Figure 2).

Table 1. Pooled user needs and preferences for healthy eating and activity.

Preferences	Total responses selected, n (%)
Eating preferences (n=122 questions deployed)	
Remind me of my daily health intention	51 (41.8)
Quick and easy option to log my food	50 (41.0)
Provide healthy eating options nearby	39 (32.0)
Show me my calorie intake for the day	37 (30.3)
Something else[a]	17 (13.9)
Video or text chat with a support person	6 (4.9)
Activity preferences (n=28 questions deployed)	
Show me my activity progress for the day	11 (39)
Remind me of my daily health intention	9 (32)
Provide easy exercise options nearby	9 (32)
Video demonstration of easy exercise	8 (29)
Something else[b]	5 (18)
Easy way to log my activity	4 (14)

[a]Access to healthy preferred options or premade healthy options; a list of 3-5 quick and healthy meal options to prepare in advance; someone else to prepare healthy meals; inventory of food in my house or assistance figuring out healthy options without getting up; reduced focus on calories; suggest healthy options based on what I have in my house; motivational messaging and a dashboard of my progress; recipe suggestions; suggestions on healthy options in my area or just general food suggestions.

[b]Someone to go with me and hold me accountable; support me by checking in with me like a friend would.

Usability, Acceptability, Feasibility, and Appropriateness

At the end of each study day, contextualized technology adaptation process questions regarding feasibility, acceptability, and appropriateness were deployed. The majority of respondents indicated that the number of daily prompts was "just right" (37/72, 51%), followed by "too low" (23/72, 32%), with "too high" (12/72, 17%) being the least frequently selected response. The majority of respondents viewed timing of prompts as "moderately helpful" (36/72, 50%), followed by "not at all helpful" (26/72, 36%), with the fewest responses indicating timing was "extremely helpful" (10/72, 14%). Responses to the final question, "What would have made this application more helpful," shown in Table 2, indicated that users preferred personalized messaging and simplified self-monitoring options. Participants also had the option of adding free-text comments, as Table 3 shows.

Table 2. Pooled data for all participants (n=19, total possible responses n=69) responding to acceptability, feasibility, and appropriateness questions deployed at the end of each day.

What would have made this application more helpful?	Total responses, n (%)
Personalized text messages about how to reach my health goals	29 (42)
Provide an easy way to log my food and activity	28 (41)
Incentive for healthy choices (eg, collecting “points” to be used toward a reward)	28 (4
Anticipate sedentary behavior and send me activity suggestions	27 (39)
Anticipate unhealthy eating and send me healthier options nearby	25 (36)
Provide interactive video of exercise examples	23 (33)
Provide healthy meal options and recipes	23 (33)
Track and show me my activity throughout the day	21 (30)
Ability to video chat with a professional	5 (7)
Link to social media to share my progress with others	1 (1)
See a friend or loved one’s progress	0 (0)

Table 3. Postparticipation interview responses vs app-collected responses.

Question	Representative interview quote	Representative app quote
What did you like about the app?	I liked setting a health goal and then checking in at the end of the day.	I know how I should be feeling and behaviors I should be changing, but it is really hard when I am all alone to take the right steps...[the app] kept me busy so I wouldn’t get so sad.
What did you dislike about the app?	Too many prompts.	Hounding people to eat healthy and exercise is the #1 way to drive an emotionally fragile person away.
What would have made the app more useful?	Make logging easier.	Maybe a mood tracker would be helpful, and a visual of my progress.
What would make you more likely to continue using an app like this?	I want to be able to change the notification settings.	I’d like to be able to ask for help making eating choices when I’m ready to eat. Due to my disability, I don’t have a regular sleep schedule so I don’t eat meals at regular times.

In order to contrast results from traditional acceptability, feasibility, and appropriateness assessment methods with data collected via the app, we conducted semistructured interviews with 4 users representing high and low user engagement, contrasted with responses provided via free text through the app (Table 3). Follow-up interviewees in general gave positive feedback about the context, particularly about the intention setting at the beginning of the day, and anticipated or actual difficulty rating features. In contrast, feedback obtained through the app on a daily basis included more detail about the usability of the needs assessment tool, particularly with respect to what participants didn’t like. Respondents noted that their day-to-day schedules varied, in particular related to sleep and mood, so prescheduled “in-the-moment” needs assessment prompts weren’t always relevant and were perceived as annoying if they came at a time that was busy or stressful. Respondents also expressed a desire for the app to provide more emotional support, learn their patterns and behaviors, and personalize goals or substitute behaviors that consider a specific condition or disability.

Discussion

Principal Findings

We tested the utility of a novel mobile technology–based technique to capture both group-based and individually dynamic user needs and preferences. We conducted this assessment, which we term ecological daily needs assessment, alongside a basic mHealth intervention designed to support health behaviors linked to healthy weight management. There were 3 key findings. First, the technique demonstrated the ability to detect dynamic needs and preferences, which changed over time differentially within an individual, as well as between individuals. The technique collected group-level usability data suggesting specific adaptations for improving the app, as well as being informative for intervention adaptation. Second, the number of daily prompts was acceptable, but timing was rarely appropriate despite being during prespecified times indicated by the participant. Response rates declined significantly after 3 days, suggesting a possible threshold for collecting useful needs assessment data. Third, this approach to usability assessment yielded specific critical feedback, contrasted with less specific and generally positive or neutral feedback obtained from the more traditional method of semistructured interviews. These results highlight the importance of context in usability testing of mobile assessments, which can yield data relevant to

dynamic needs and preferences at the group and individual levels.

Digital tools are commonly leveraged to temporally link active, passive, and metadata to improve user engagement. Despite this, mHealth apps are still disabled within weeks of download [38], suggesting that extensive digital phenotyping efforts are insufficient for optimizing user engagement [39,40]. Assessment of user preferences increases mHealth treatment engagement [41], but populations perceived as difficult to engage, such as individuals with psychiatric illnesses, are often excluded from treatment adaptation research, primarily due to concerns about the reliability of self-report measures due to cognitive limitations [42]. However, it is exactly these cognitive limitations that necessitate the need for usability and preference testing [41,43]. The extant research evaluating mHealth or digital weight loss treatments in young adults with mental illness is limited, but suggests that the unique needs and preferences of this population may have important implications for mHealth treatment development [44]. These results underscore the importance of ecologically valid data in developing interventions that effectively engage this population [45].

The needs assessment in this study was anchored within a behavior change framework based on existing treatments [29,46] to provide the user exposure to an intervention upon which they could consider their needs and preferences. In contrast, more traditional usability assessments conducted at the end of an intervention, even when paired with utilization data, are subject to recall bias and do not capture how end users experience the app in real time [47]. The complexity of programming needed to build a learning system from inception contributes to cost, and there are as yet few cost-benefit data to inform investment in artificial intelligence features like machine learning for health apps [48]. Developers of mHealth treatments in resource-constrained settings might more expeditiously begin app development with an anchored needs assessment in place to better understand end user needs, which could provide useful information regarding which data inputs are most relevant to employ in learning algorithms.

Limitations

It is important to note limitations of this study. First, the number of participants was small, but within the recommended range for usability testing. Second, participation was limited to individuals owning iOS devices. Apps on non-iOS smart devices using the Android operating system, which are more commonly used in psychiatric populations [49], are needed. Third, although the needs assessment technique in this study prioritized participant-scheduled prompts, participants were not able to change the schedule or initiate a response when they encountered a need outside of the schedule, resulting in potential missed opportunities to capture more nuanced aspects of individual variability in needs and preferences. Fourth, important questions remain regarding the validity and relevance of our observations for mHealth treatment development. If individually dynamic needs and preferences are relevant, what is the best way to accurately capture this information? And how can this information be meaningfully incorporated into behavioral mHealth intervention development and adaptation? Empirical testing is needed to determine whether capturing individually dynamic data leads to superior mHealth interventions and outcomes.

Conclusions

The results of this study provide insights that may inform the development of self-adapting treatments, which this study and others have identified as particularly germane to engagement of individuals with chronic health concerns, including mental health conditions [27,50]. In populations where app engagement is linked to treatment outcome, or in settings where funding for development is limited, early usability testing of digital features, using the digital context, is a low-cost option for determining which aspects of development to prioritize with limited funding. The resulting contextually relevant information might be particularly useful in guiding real-time treatment adaptation while limiting in-person contact, which will likely be important for the future of clinical research in vulnerable populations during public health events like the coronavirus pandemic [51]. Additional study is needed to determine whether a mobile needs assessment can usefully inform behavioral treatment development for diverse patient populations and operating systems.

Acknowledgments

This research was made possible by funding from the Center for Brain Research in Mood Disorders and the Healthy Mind Lab in the Department of Psychiatry at Washington University School of Medicine, St. Louis, MO, USA, and the US National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; R01DK103760). In addition, AT received grant support from the National Institute on Drug Abuse. This publication was made possible by grant no. 1P30DK092950 from the NIDDK, and its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIDDK. The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding and senior author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Conflicts of Interest

GEN reports grant support from the US National Institute of Mental Health, the Barnes Jewish Hospital Foundation, the Center for Diabetes Translational Research, and the McDonnell Center for Systems Neuroscience at Washington University, and has consulted for Alkermes and Sunovion. CLM is the developer of the Status/Post software used to create the mobile app reported

on in this paper. EJL reports grant support from Janssen, Aptinyx, Acadia, Lundeck, Takeda, PCORI, the Barnes Jewish Hospital Foundation, and the McKnight Brain Research Foundation, and has consulted for Janssen and Jazz Pharmaceuticals.
CLM is the developer of the Status/Post software used to create the mobile app reported on in this paper.
EJL reports grant support from Janssen, Aptinyx, Acadia, Lundeck, Takeda, PCORI, the Barnes Jewish Hospital Foundation, and the McKnight Brain Research Foundation, and has consulted for Janssen and Jazz Pharmaceuticals.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 17268 KB - mhealth_v8i11e18609_app1.docx]

References

1. Buttorff C, Ruder T, Bauman M. Multiple Chronic Conditions in the United States. Santa Monica, CA: RAND Corporation; 2017. URL: https://www.rand.org/pubs/tools/TL221.html [accessed 2020-10-19]
2. NHE Fact Sheet. Baltimore, MD: U.S. Centers for Medicare & Medicaid Services; 2020 Mar 24. URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet [accessed 2020-10-19]
3. 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]
4. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
5. Bowling A, Ebrahim S. Measuring patients' preferences for treatment and perceptions of risk. Qual Health Care 2001 Sep;10 Suppl 1:i2-i8. [doi: 10.1136/qhc.0100002] [Medline: 11533430]
6. Franz MJ, VanWormer JJ, Crain AL, Boucher JL, Histon T, Caplan W, et al. Weight-loss outcomes: a systematic review and meta-analysis of weight-loss clinical trials with a minimum 1-year follow-up. J Am Diet Assoc 2007 Oct;107(10):1755-1767. [doi: 10.1016/j.jada.2007.07.017] [Medline: 17904936]
7. Dombrowski SU, Knittle K, Avenell A, Araújo-Soares V, Sniehotta FF. Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. BMJ 2014;348:g2646. [Medline: 25134100]
8. Byrne S, Cooper Z, Fairburn C. Weight maintenance and relapse in obesity: a qualitative study. Int J Obes Relat Metab Disord 2003 Aug;27(8):955-962. [doi: 10.1038/sj.ijo.0802305] [Medline: 12861237]
9. Yu Z, Sealey-Potts C, Rodriguez J. Dietary self-monitoring in weight management: current evidence on efficacy and adherence. J Acad Nutr Diet 2015 Dec;115(12):1931-1938. [doi: 10.1016/j.jand.2015.04.005] [Medline: 26028176]
10. Johnston CA, Tyler C, Foreyt JP. Behavioral management of obesity. Curr Atheroscler Rep 2007 Dec;9(6):448-453. [doi: 10.1007/s11883-007-0060-1] [Medline: 18377784]
11. Melamed OC, Fernando I, Soklaridis S, Hahn MK, LeMessurier KW, Taylor VH. Understanding engagement with a physical health service: a qualitative study of patients with severe mental illness. Can J Psychiatry 2019 Dec;64(12):872-880 [FREE Full text] [doi: 10.1177/0706743719862980] [Medline: 31303027]
12. Firth J, Rosenbaum S, Stubbs B, Gorczynski P, Yung AR, Vancampfort D. Motivating factors and barriers towards exercise in severe mental illness: a systematic review and meta-analysis. Psychol Med 2016 Dec;46(14):2869-2881 [FREE Full text] [doi: 10.1017/S0033291716001732] [Medline: 27502153]
13. Nicol G, Worsham E, Haire-Joshu D, Duncan A, Schweiger J, Yingling M, et al. Getting to more effective weight management in antipsychotic-treated youth: a survey of barriers and preferences. Child Obes 2016 Feb;12(1):70-76 [FREE Full text] [doi: 10.1089/chi.2015.0076] [Medline: 26788619]
14. Lenhart A. Washington, DC: Pew Research Center; 2015 Apr 09. URL: https://www.pewresearch.org/internet/2015/04/09/teens-social-media-technology-2015/ [accessed 2020-10-19]
15. Jacobs S, Radnitz C, Hildebrandt T. Adherence as a predictor of weight loss in a commonly used smartphone application. Obes Res Clin Pract 2017;11(2):206-214. [doi: 10.1016/j.orcp.2016.05.001] [Medline: 27292942]
16. Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. Am J Prev Med 2008 Aug;35(2):177-181 [FREE Full text] [doi: 10.1016/j.amepre.2008.05.001] [Medline: 18550322]
17. De Hert M, Correll CU, Bobes J, Cetkovich-Bakmas M, Cohen D, Asai I, et al. Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. World Psychiatry 2011 Feb;10(1):52-77. [doi: 10.1002/j.2051-5545.2011.tb00014.x] [Medline: 21379357]
18. Wiltsey Stirman S, Baumann AA, Miller CJ. The FRAME: an expanded framework for reporting adaptations and modifications to evidence-based interventions. Implement Sci 2019 Jun 06;14(1):58 [FREE Full text] [doi: 10.1186/s13012-019-0898-y] [Medline: 31171014]

19. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251. [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
20. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
21. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [Medline: 18509902]
22. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. Schizophr Bull 2014 Nov;40(6):1244-1253. [doi: 10.1093/schbul/sbu033] [Medline: 24609454]
23. Huang Y, Benford S, Blake H. Digital interventions to reduce sedentary behaviors of office workers: scoping review. J Med Internet Res 2019 Feb 07;21(2):e11079 [FREE Full text] [doi: 10.2196/11079] [Medline: 30730294]
24. Fisher AJ. Toward a dynamic model of psychological assessment: implications for personalized care. J Consult Clin Psychol 2015 Aug;83(4):825-836. [doi: 10.1037/ccp0000026] [Medline: 26009781]
25. Moore RC, Depp CA, Wetherell JL, Lenze EJ. Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. J Psychiatr Res 2016 Apr;75:116-123. [doi: 10.1016/j.jpsychires.2016.01.011] [Medline: 26851494]
26. Floch J, Zettl A, Fricke L, Weisser T, Grut L, Vilarinho T, et al. User needs in the development of a health app ecosystem for self-management of cystic fibrosis: user-centered development approach. JMIR Mhealth Uhealth 2018 May 08;6(5):e113 [FREE Full text] [doi: 10.2196/mhealth.8236] [Medline: 29739742]
27. Setiawan IMA, Zhou L, Alfikri Z, Saptono A, Fairman AD, Dicianno BE, et al. An adaptive mobile health system to support self-management for persons with chronic conditions and disabilities: usability and feasibility studies. JMIR Form Res 2019 Apr 25;3(2):e12982 [FREE Full text] [doi: 10.2196/12982] [Medline: 31021324]
28. Hardeman W, Houghton J, Lane K, Jones A, Naughton F. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. Int J Behav Nutr Phys Act 2019 Apr 03;16(1):31 [FREE Full text] [doi: 10.1186/s12966-019-0792-7] [Medline: 30943983]
29. Tabak RG, Strickland JR, Stein RI, Dart H, Colditz GA, Kirk B, et al. Development of a scalable weight loss intervention for low-income workers through adaptation of interactive obesity treatment approach (iOTA). BMC Public Health 2018 Nov 16;18(1):1265 [FREE Full text] [doi: 10.1186/s12889-018-6176-0] [Medline: 30445939]
30. Stein RI, Strickland JR, Tabak RG, Dale AM, Colditz GA, Evanoff BA. Design of a randomized trial testing a multi-level weight-control intervention to reduce obesity and related health conditions in low-income workers. Contemp Clin Trials 2019 Apr;79:89-97 [FREE Full text] [doi: 10.1016/j.cct.2019.01.011] [Medline: 30664943]
31. Degroote L, DeSmet A, De Bourdeaudhuij I, Van Dyck D, Crombez G. Content validity and methodological considerations in ecological momentary assessment studies on physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2020 Mar 10;17(1):35 [FREE Full text] [doi: 10.1186/s12966-020-00932-9] [Medline: 32151251]
32. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
33. Sweller J. Cognitive load theory, learning difficulty, and instructional design. Learn Instruct 1994 Jan;4(4):295-312. [doi: 10.1016/0959-4752(94)90003-5]
34. Ahmad FA, Payne PRO, Lackey I, Komeshak R, Kenney K, Magnusen B, et al. Using REDCap and Apple ResearchKit to integrate patient questionnaires and clinical decision support into the electronic health record to improve sexually transmitted infection testing in the emergency department. J Am Med Inform Assoc 2020 Feb 01;27(2):265-273. [doi: 10.1093/jamia/ocz182] [Medline: 31722414]
35. Lyon AR, Wasse JK, Ludwig K, Zachry M, Bruns EJ, Unützer J, et al. The contextualized technology adaptation process (CTAP): optimizing health information technology to improve mental health systems. Adm Policy Ment Health 2016 May;43(3):394-409. [doi: 10.1007/s10488-015-0637-x] [Medline: 25677251]
36. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. New York, NY: Association for Computing Machinery; 1993 Presented at: INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems; Apr 24-29, 1993; Amsterdam, Netherlands p. 24-213. [doi: 10.1145/169059.169166]
37. Macefield R. How to specify the participant group size for usability studies: a practitioner's guide. J Usabil Stud 2009;5(1):34-45.
38. Dorsey ER, Yvonne CY, McConnell MV, Shaw SY, Trister AD, Friend SH. The use of smartphones for health research. Acad Med 2017 Feb;92(2):157-160. [doi: 10.1097/ACM.0000000000001205] [Medline: 27119325]
39. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015 May 12;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
40. Insel TR. Digital phenotyping: technology for a new science of behavior. JAMA 2017 Oct 03;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]

41. Søgaard Neilsen A, Wilson RL. Combining e-mental health intervention development with human computer interaction (HCI) design to enhance technology-facilitated recovery for people with depression and/or anxiety conditions: an integrative literature review. Int J Ment Health Nurs 2019 Feb;28(1):22-39. [doi: 10.1111/inm.12527] [Medline: 30133096]
42. Morrison LG, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. Telemed J E Health 2012 Mar;18(2):137-144. [doi: 10.1089/tmj.2011.0062] [Medline: 22381060]
43. Archer N, Keshavjee K, Demers C, Lee R. Online self-management interventions for chronically ill patients: cognitive impairment and technology issues. Int J Med Inform 2014 Apr;83(4):264-272. [doi: 10.1016/j.ijmedinf.2014.01.005] [Medline: 24507762]
44. Aschbrenner KA, Naslund JA, Shevenell M, Mueser KT, Bartels SJ. Feasibility of behavioral weight loss treatment enhanced with peer support and mobile health technology for individuals with serious mental illness. Psychiatr Q 2016 Sep;87(3):401-415 [FREE Full text] [doi: 10.1007/s11126-015-9395-x] [Medline: 26462674]
45. Naslund JA, Aschbrenner KA. Digital technology for health promotion: opportunities to address excess mortality in persons living with severe mental disorders. Evid Based Ment Health 2019 Feb;22(1):17-22 [FREE Full text] [doi: 10.1136/ebmental-2018-300034] [Medline: 30559332]
46. Nicol GE, Kolko R, Lenze EJ, Yingling MD, Miller JP, Ricchio AR, et al. Adiposity, hepatic triglyceride, and carotid intima media thickness during behavioral weight loss treatment in antipsychotic-treated youth: a randomized pilot study. J Child Adolesc Psychopharmacol 2019 Aug;29(6):439-447 [FREE Full text] [doi: 10.1089/cap.2018.0120] [Medline: 30994376]
47. Chen AT, Wu S, Tomasino KN, Lattie EG, Mohr DC. A multi-faceted approach to characterizing user behavior and experience in a digital mental health intervention. J Biomed Inform 2019 Jun;94:103187 [FREE Full text] [doi: 10.1016/j.jbi.2019.103187] [Medline: 31026595]
48. Wolff J, Pauling J, Keck A, Baumbach J. The economic impact of artificial intelligence in health care: systematic review. J Med Internet Res 2020 Feb 20;22(2):e16866 [FREE Full text] [doi: 10.2196/16866] [Medline: 32130134]
49. Young AS, Cohen AN, Niv N, Nowlin-Finch N, Oberman RS, Olmos-Ochoa TT, et al. Mobile phone and smartphone use by people with serious mental illness. Psychiatr Serv 2020 Mar 01;71(3):280-283. [doi: 10.1176/appi.ps.201900203] [Medline: 31744429]
50. Thornton LK, Kay-Lambkin FJ. Specific features of current and emerging mobile health apps: user views among people with and without mental health problems. Mhealth 2018;4:56 [FREE Full text] [doi: 10.21037/mhealth.2018.11.04] [Medline: 30701174]
51. Padala PR, Jendro AM, Padala KP. Conducting clinical research during the COVID-19 pandemic: investigator and participant perspectives. JMIR Public Health Surveill 2020 Apr 06;6(2):e18887 [FREE Full text] [doi: 10.2196/18887] [Medline: 32250281]

Abbreviations

mHealth: mobile health

Edited by G Eysenbach, L Buis; submitted 10.03.20; peer-reviewed by A Chen, DR Wahl, A Takano; comments to author 20.04.20; revised version received 14.06.20; accepted 15.10.20; published 13.11.20.

Please cite as:
Nicol GE, Ricchio AR, Metts CL, Yingling MD, Ramsey AT, Schweiger JA, Miller JP, Lenze EJ
A Smartphone-Based Technique to Detect Dynamic User Preferences for Tailoring Behavioral Interventions: Observational Utility Study of Ecological Daily Needs Assessment
JMIR Mhealth Uhealth 2020;8(11):e18609
URL: http://mhealth.jmir.org/2020/11/e18609/
doi:10.2196/18609
PMID:33055063

Original Paper

Evaluation of the MoMba Live Long Remote Smoking Detection System During and After Pregnancy: Development and Usability Study

Stephanie Valencia[1,2], BSc; Laura Callinan[3], MPH; Frederick Shic[1,4], PhD; Megan Smith[3], MPH, DrPH

[1]Child Study Center, Yale University School of Medicine, New Haven, CT, United States

[2]Carnegie Mellon University, Pittsburgh, PA, United States

[3]Department of Psychiatry, Yale University School of Medicine, New Haven, CT, United States

[4]Seattle's Children's Research Institute, University of Washington School of Medicine, Seattle, WA, United States

Corresponding Author:
Megan Smith, MPH, DrPH
Department of Psychiatry
Yale University School of Medicine
230 South Frontage Road
New Haven, CT, 06519
United States
Phone: 1 203 764 8655
Email: megan.smith@yale.edu

Abstract

Background: The smoking relapse rate during the first 12 months after pregnancy is around 80% in the United States. Delivering remote smoking cessation interventions to women in the postpartum period can reduce the burden associated with frequent office visits and can enable remote communication and support. Developing reliable, remote, smoking measuring instruments is a crucial step in achieving this vision.

Objective: The study presents the evaluation of the MoMba Live Long system, a smartphone-based breath carbon monoxide (CO) meter and a custom iOS smartphone app. We report on how our smoking detection system worked in a controlled office environment and in an out-of-office environment to examine its potential to deliver a remote contingency management intervention.

Methods: In-office breath tests were completed using both the MoMba Live Long system and a commercial monitor, the piCO$^+$ Smokerlyzer. In addition, each participant provided a urine test for smoking status validation through cotinine. We used in-office test data to verify the validity of the MoMba Live Long smoking detection system. We also collected out-of-office tests to assess how the system worked remotely and enabled user verification. Pregnant adult women in their second or third trimester participated in the study for a period of 12 weeks. This study was carried out in the United States.

Results: Analyses of in-office tests included 143 breath tests contributed from 10 participants. CO readings between the MoMba Live Long system and the piCO$^+$ were highly correlated (r=.94). In addition, the MoMba Live Long system accurately distinguished smokers from nonsmokers with a sensitivity of 0.91 and a specificity of 0.94 when the piCO$^+$ was used as a gold standard, and a sensitivity of 0.81 and specificity of 1.0 when cotinine in urine was used to confirm smoking status. All participants indicated that the system was easy to use.

Conclusions: Relatively inexpensive portable and internet-connected CO monitors can enable remote smoking status detection in a wide variety of nonclinical settings with reliable and valid measures comparable to a commercially available CO monitor.

Trial Registration: ClinicalTrials.gov NCT02237898; https://clinicaltrials.gov/ct2/show/NCT02237898

(JMIR Mhealth Uhealth 2020;8(11):e18809) doi:10.2196/18809

KEYWORDS

breath carbon monoxide; contingency management; smoking cessation; pregnancy; mobile-based sensor; mobile phone

Introduction

Background

Tobacco is the most common substance of abuse used during pregnancy, with a substantially higher rate of use among socioeconomically disadvantaged women [1]. While approximately 50% of women who previously smoked regularly abstain from smoking in pregnancy, relapse to smoking postnatally remains a challenge, with over 70% of women who remitted relapsing within 12 months postpartum in the United States [2,3].

Contingency management (CM), rewarding financial incentives contingent upon biochemically verified abstinence from recent smoking, has consistently been shown to decrease the use of tobacco [2-4] and to be an efficacious intervention for promoting smoking cessation in pregnant and postpartum women [5]. Delivering CM interventions typically involves frequent monitoring of smoking status with daily or weekly office visits. In this work, we develop a remote smoking assessment system to more widely disseminate smoking cessation interventions into community settings and to reach overburdened and underserved populations of smokers, such as expectant mothers of lower socioeconomic status.

Technologies for Remote Smoking Assessments

Mobile technologies, especially web-enabled or smartphone technology, can be used to access real-world community settings, as the access to smartphones has increased: 81% of Americans owned a smartphone in 2019 [6]. Mobile health solutions are emerging, creating an opportunity for the expansion of evidence-based practices [7-9]. In combination with sensors that enable remote biochemical assessment of smoking status, such as breath carbon monoxide (CO) measurements [10-12], mobile technologies could allow remote delivery of smoking cessation interventions.

A randomized, controlled, parallel-group design study carried out a 6-week intervention evaluating the effectiveness of an internet-based smoking cessation program in the United States using a system called Motiv8 [13]. Participants used Motiv8 to submit videos and values of breath CO tests taken with a commercial CO breath analyzer, the piCO$^+$ Smokerlyzer (Bedfont Scientific), and to confirm identity and veracity of tests through a secured website. Even with some limitations (eg, the need for a desktop computer), the study demonstrated that participants receiving rewards based on abstinence were more likely to post negative CO samples on the website than the participants who received monetary rewards independent of smoking status (odds ratio 4.56, 95% CI 2.10-9.52). More recently, breath CO manufacturers have developed commercial CO monitors that leverage the advantages of internet-based systems with new, smaller CO monitors such as the iCO Smokerlyzer (Bedfont Scientific). The iCO Smokerlyzer connects to a smartphone for personal smoking behavior monitoring and enables users to measure their CO levels but does not include a way to verify that the intended user completed the breath test appropriately. This verification is important for remote CM interventions, where additional evidence is needed to verify test validity as there are no present observers. Recent work has investigated unsupervised breath test validity by verifying exhalation through the use of pressure sensors and the use of facial recognition for user authentication in the context of smoking and alcohol use [14-16].

The MoMba Live Long System

Overview

The MoMba Live Long system consisted of a custom and portable breath CO meter that wirelessly interfaced via Bluetooth with an iOS-only app (see Figure 1). The MoMba Live Long iOS app provided an interface for the breath CO meter that allowed the participant to receive notifications regarding the availability of the breath test, taking the breath test, seeing results of smoking status, and keeping track of rewards and progress. In addition, the app verified that the participant was correctly taking the breath test by recording pictures using the front-facing camera as well as recording audio to verify that the participant was exhaling while taking the test. When a breath test indicated smoking abstinence, the participant was rewarded with tokens that could be exchanged for gift cards. The MoMba Live Long app also enabled the delivery of questionnaires through the app. The MoMba Live Long system was based on a successful app design developed to support maternal mental health in the postpartum period [17].

Figure 1. The custom and portable breath carbon monoxide (CO) meter is comprised of an environmental CO gas sensor and a custom 3D-printed chassis. A smartphone was loaded with a custom app (MoMba Live Long) that enables participants to complete scheduled breath tests and collect rewards. The Sensordrone and chassis sizes are shown at scale with respect to the phone in the image.

MoMba Breath CO Meter

The custom breath CO meter used an electrochemical CO gas sensor—the Sensordrone, part No. SDRONEG1 (Sensorcon Inc)—designed to measure environmental CO. The Sensordrone was accurate to within 10% with a resolution of 1 parts per million (ppm), making it comparable to existing commercial medical-grade CO breath analyzers [18]. We designed a 3D-printed chassis to encase the Sensordrone and allow proper air flow. 3D printing was selected as a fabrication method that allowed us to prototype in a timely manner and test different designs. The 3D-printed chassis held an inline activated-carbon filter and cotton that removed non-CO gases and excess moisture.

Breath Sample Collection

To determine smoking status, a user completed a breath test by blowing into our custom breath CO meter composed of a disposable mouthpiece connected to the custom 3D-printed chassis encasing the Sensordrone. The breath CO meter measured CO concentration in the breath after collecting exhaled air samples for 20 seconds at a 5-Hz sampling rate. Collected data were securely sent and stored in a back-end server monitored by staff.

The MoMba CO Estimate

A previous study we conducted established that the most accurate prediction of smoking status from breath was by sampling the portion of breath at the end of exhalation, which represented the middle portion of an exhalation of 20 seconds, on average [19]. During this initial testing, we detected baseline reading offsets attributed to increased temperature in the Sensordrone from recent charging and environmental pollution. This led to the addition of a baseline correction to our algorithm. Our previous findings presented an exhaustive exploration of different smoking detection approaches using different classification models and features, while this report presents an evaluation of our system's performance with our implemented algorithm that used the best-performing features from our previous study [19]. In order to obtain one single value from a 20-second sample of breath tests, we performed the following steps: (1) calculated the baseline reading for the Sensordrone, (2) located the general maximum and captured values above 50% of the maximum value, and (3) calculated the median value and removed any offsets detected in the baseline value. This final value represented the MoMba CO estimate.

Aim of the Study

We present the design of the MoMba Live Long , mobile, breath CO meter and a pilot evaluation of the feasibility of the system as a smoking assessment tool during and after pregnancy.

Methods

Participants

The MoMba Live Long pilot study was approved by the ethical review board of the Yale School of Medicine. Participants were recruited at local clinics; through community outreach, flyers, and advertisements; and from referrals. Participants were women who met the following inclusion criteria: were 18 years of age or older, were daily tobacco cigarette smokers not using nicotine replacement therapy, were pregnant with a singleton in their second or third trimester, and had a desire to stop smoking. Women were not eligible if English was not their primary language, if they did not live in the city in which the study was conducted, if they did not plan to deliver their baby at the local

hospital, or if they met medical exclusions, as determined by medical record review, including respiratory medical conditions such as chronic obstructive pulmonary disease, HELLP (hemolysis, elevated liver enzymes, low platelet count) syndrome, or pregnancy complications. If a participant lost her child, she was no longer eligible to participate. Eligible participants completed an intake visit where informed written consent was provided and smoking status was confirmed with a urine sample.

In-Office and Remote Data Collection Procedures

During the MoMba Live Long pilot study, breath tests were collected for 12 weeks and followed the CM schedule described by Higgins et al [20]. For each consecutive smoke-free breath test, a participant received a higher number of tokens per test. Breath samples were obtained via a mixture of in-office and remote tests. Each participant was asked to attend up to 18 in-office visits: five visits in week 1, two visits in weeks 2 and 3, and one visit per week for weeks 4-12. In-office visits did not always coincide with a CM-related breath test. Before completing breath tests, participants were asked to report the number of cigarettes smoked within the previous 24 hours and the date and time of the last cigarette smoked through the MoMba Live Long app. In-office visits consisted of first completing one breath test using the MoMba Live Long system and then one breath test with piCO$^+$. Participants also provided a urine sample to validate smoking.

During in-office testing, participants were asked to sit in a chair in an upright and straight position to accurately complete a breath test. Participants were instructed to inhale deeply and hold their breath for 15 seconds and then exhale all the air completely (up to 20 seconds). Holding the breath allowed the sensor to obtain a CO value close to alveolar CO concentration [21]. Both the MoMba Live Long system and the piCO$^+$ monitor displayed a countdown on their respective screens as participants were holding their breath. Once the countdown was finished, the participant blew into the sensor. Both systems displayed a second countdown to indicate that the sensor was collecting the sample.

All remote tests followed the same procedures as in-office breath tests. The participants received a kit to take home that included an iPhone, which participants were encouraged to use as their primary phone; a chassis; a Sensordrone; a charger for the Sensordrone; disposable mouthpieces; and replacement filters. After participants completed their tests, research staff verified the validity of the test using the back-end server and approved the corresponding financial rewards. Participants received a notification to complete their test between the hours of 8:30 AM and 3 PM EST. The notification for the breath test expired after 5 hours.

Measures

Participant Characteristics at Intake

A baseline questionnaire asked questions regarding demographics and smoking habits. The Fagerstrom Test for Cigarette Dependence was used to assess dependence on nicotine; a higher score indicates greater dependence [22].

MoMba Performance Against Gold-Standard Measures

The primary outcome measure for this study was the validity of the MoMba CO outcome; this measure was compared with the gold standard of piCO$^+$ and urine cotinine collected at the same visit. A measure of smoking abstinence was defined as CO-negative breath samples determined by the MoMba breath CO meter and the piCO$^+$ (≤6 ppm). The selected CO cutoff levels were recommended by previous work [10,23]. All urine tests were tested for adulteration using a specimen validity test and then sent to a lab for a quantitative cotinine urine assay with quality control. A cotinine concentration value less than 50 ng/mL indicated smoking abstinence [24]. Due to limits in lab detection, cotinine concentrations reported as less than 10 ng/mL were estimated as 5 ng/mL.

Variation of CO Values Since Last Cigarette Smoked

To examine how breath CO values varied since last cigarette smoked, the half-life of breath CO, 3-6 hours [12], was used to create a dichotomous variable indicating if the participant smoked within 5 hours of a breath test or more than 5 hours before a breath test.

Variation of CO Values According to Pregnancy Status

Pregnancy and postpregnancy CO values were compared to explore any potential differences; postpregnancy status was determined by a participant's delivery date. We used the number of cigarettes and time since last cigarette smoked to investigate if these variables impacted the observed pattern of results.

Out-of-Office Performance and User Experience

A follow-up questionnaire was asked 3 months after intake, at the completion of CM; this questionnaire included two Likert scale questions regarding how easy it was to use (1) the sensor and (2) the MoMba Live Long app; response options were 1 (Extremely Difficult), 2 (Difficult), 3 (Neutral), 4 (Easy), and 5 (Extremely Easy). In addition, we evaluated the completion and delivery of the remote breath test as well as the performance of the front-facing camera authentication method and the microphone to detect exhalation during remote breath tests.

Statistical Analysis

Categorical variables were described with count and proportions. Normally distributed data were presented with mean and SD; nonnormally distributed data were presented with median and IQR. Sensitivity was defined as the proportion of tests for which a true-positive breath CO test was detected. Specificity was defined as the proportion of tests for which there was a true-negative breath CO test. Receiver operating characteristic curves were generated for the MoMba CO estimate and the piCO$^+$ by plotting the percentage of true positives against the percentage of false positives. Area under the curve (AUC) was calculated for each plot using the pROC package for R (The R Foundation) [25]. Since each participant contributed multiple observations, methods to calculate sensitivity and specificity accounted for clustering due to participant [26]. To assess the relationship across smoking measures, a correlation coefficient was calculated accounting for repeated observations of participants [27]. Generalized linear models were used to compare smoking status indicators between groups, accounting

for potentially correlated data among participants. An autoregressive correlation structure was specified for the models. Time since last cigarette smoked and/or number of cigarettes smoked in the past 24 hours were used as covariates in the model. Statistical significance was determined as $P<.05$ (2-tailed).

Results

Participant Characteristics at Intake

A total of 10 pregnant adult women participated in the MoMba Live Long pilot. The majority of women were single and never married (7/10, 70%); 50% (5/10) of the women were Black or African American, non-Hispanic (see Table 1). The average age of participants was 31.7 years (SD 4.6).

Table 1. Demographics and clinical characteristics of participants from the MoMba Live Long pilot study.

Characteristic	Value (N=10)
Marital status, n (%)	
Single and never married	7 (70)
Married or partnered	3 (30)
Race and ethnicity, n (%)	
White, non-Hispanic	4 (40)
Black or African American, non-Hispanic	5 (50)
White, Hispanic	1 (10)
Highest year of education completed, n (%)	
Grades 9-12 or General Educational Diploma	5 (50)
At least 1 year of college or vocational school	5 (50)
How many cigarettes per day do you smoke? n (%)	
0-5 cigarettes	0 (0)
6-10 cigarettes	9 (90)
11-20 cigarettes	1 (10)
Age in years, mean (SD)	31.7 (4.6)
Previous pregnancies not including current pregnancy, mean (SD)	3.7 (2.5)
Number of weeks pregnant, mean (SD)	25.9 (9.1)
Fagerstrom score[a], mean (SD)	4.5 (1.6)

[a]Fagerstrom Test of Cigarette Dependence conducted at intake; scores range from 1 to 10.

MoMba Performance Against Gold-Standard Measures

A total of 143 breath tests were collected; participants contributed, on average, 14.3 (SD 3.5) tests (range 8-17). Significant CO reading correlations (r=.94) were observed between the MoMba CO estimate and the piCO$^+$ (see Table 2). When using the piCO$^+$ as the gold standard, the MoMba CO estimate presented a sensitivity of 0.91 and specificity of 0.94.

Of the 139 tests with urine data, a moderate linear relationship was seen between both CO breath measures and the urine cotinine tests (see Table 2): MoMba CO estimate (r=.52) and piCO$^+$ (r=.57). When using cotinine in urine as the gold-standard smoking indicator, the sensitivity with the MoMba CO estimate was 0.81 and the piCO$^+$ measure showed a sensitivity of 0.87. Both MoMba CO and piCO$^+$ estimates had a specificity of 1.0.

The AUC for the MoMba CO estimate was 0.95 (95% CI 0.91-0.99) when the piCO$^+$ measure was used as a gold standard (see Figure 2). With urine cotinine as the gold standard, the AUC for the MoMba CO estimate was 0.95 (95% CI 0.92-0.99) and for the piCO$^+$ measure was 0.99 (95% CI 0.99-1.0) (see Figure 3).

Table 2. Sensitivity and specificity of the MoMba Live Long system carbon monoxide (CO) measure with the piCO$^+$ measure as gold standard, and MoMba Live Long system CO and piCO$^+$ measures with cotinine in urine as gold standard.

Smoking status test	Sensitivity (95% CI)	Specificity (95% CI)	*r*	*P* value
piCO$^+$ as gold standard				
MoMba CO estimate (n=143)	0.91 (0.80-1.0)	0.94 (0.83-1.0)	.94	<.001
Cotinine in urine as gold standard				
MoMba CO estimate (n=139)	0.81 (0.65-0.97)	1.0 (1.0-1.0)	.52	.01
piCO$^+$ (n=139)	0.87 (0.73-1.0)	1.0 (1.0-1.0)	.57	.02

Figure 2. Receiver operating characteristic curves for the MoMba Live Long system carbon monoxide (CO) measure with the piCO$^+$ measure as gold standard.

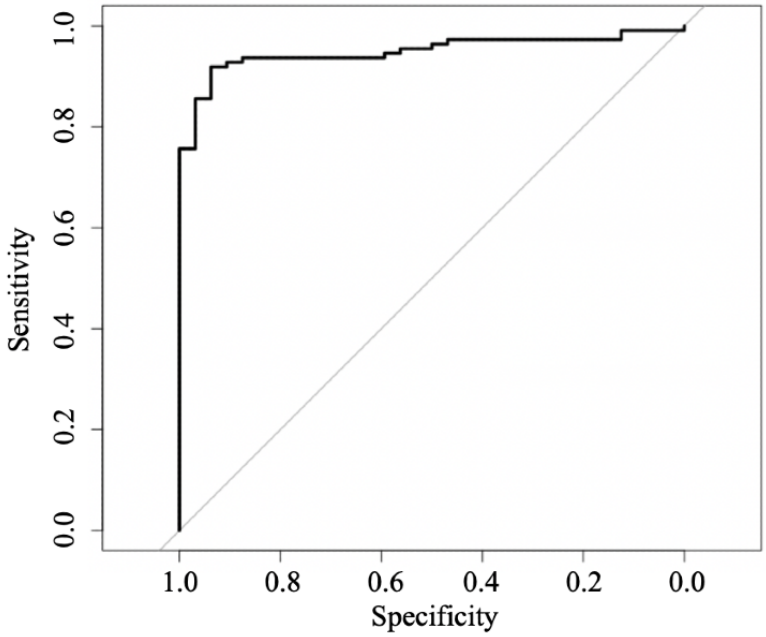
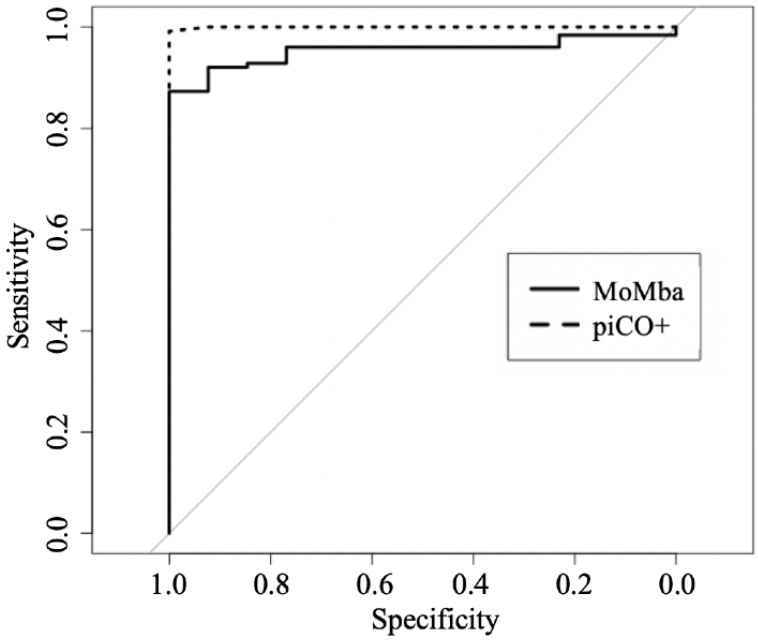


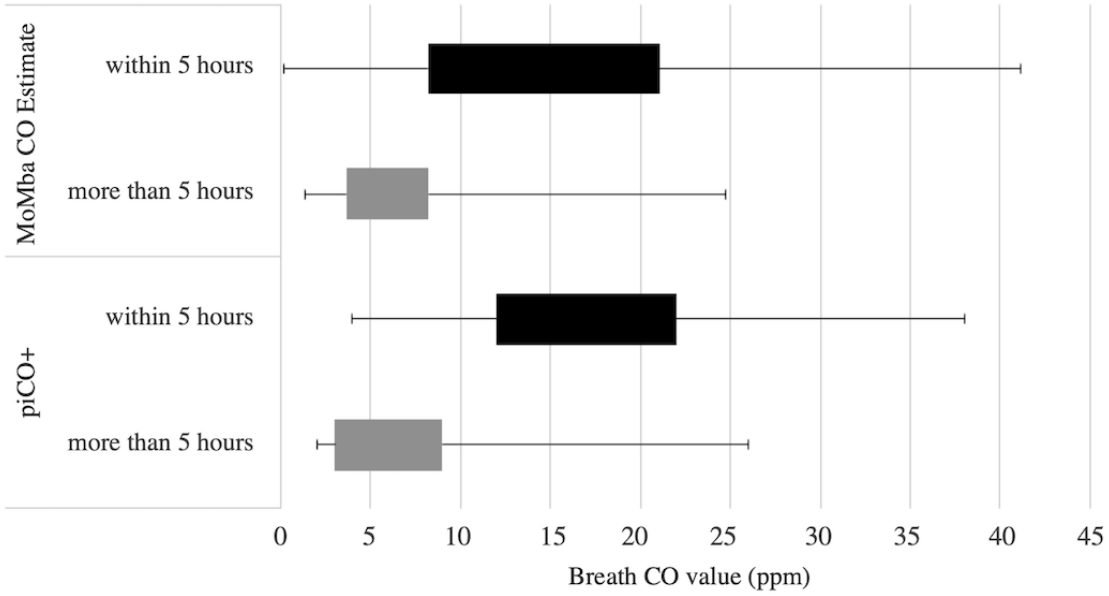
Figure 3. Receiver operating characteristic curves for the MoMba Live Long system carbon monoxide (CO) and piCO$^+$ measures with cotinine in urine as gold standard.



Variation of CO Values Since Last Cigarette Smoked

Figure 4 shows the difference in CO measures based on the dichotomized variable *time of last cigarette smoked*. Of the 143 completed in-office breath tests, 8 (5.6%) tests did not have data regarding time of last cigarette smoked, 102 (71.3%) breath tests were taken within 5 hours of smoking, and 33 (23.1%) tests were taken after 5 hours of smoking. All breath CO values were higher when a participant smoked within 5 hours of the breath test (MoMba CO estimate, P=.048; piCO$^+$, P=.03). After controlling for the number of cigarettes smoked in the past 24 hours, values remained higher for participants who smoked within 5 hours of the test compared with participants who smoked more than 5 hours before the test (MoMba CO estimate, P=.045; piCO$^+$, P=.02).

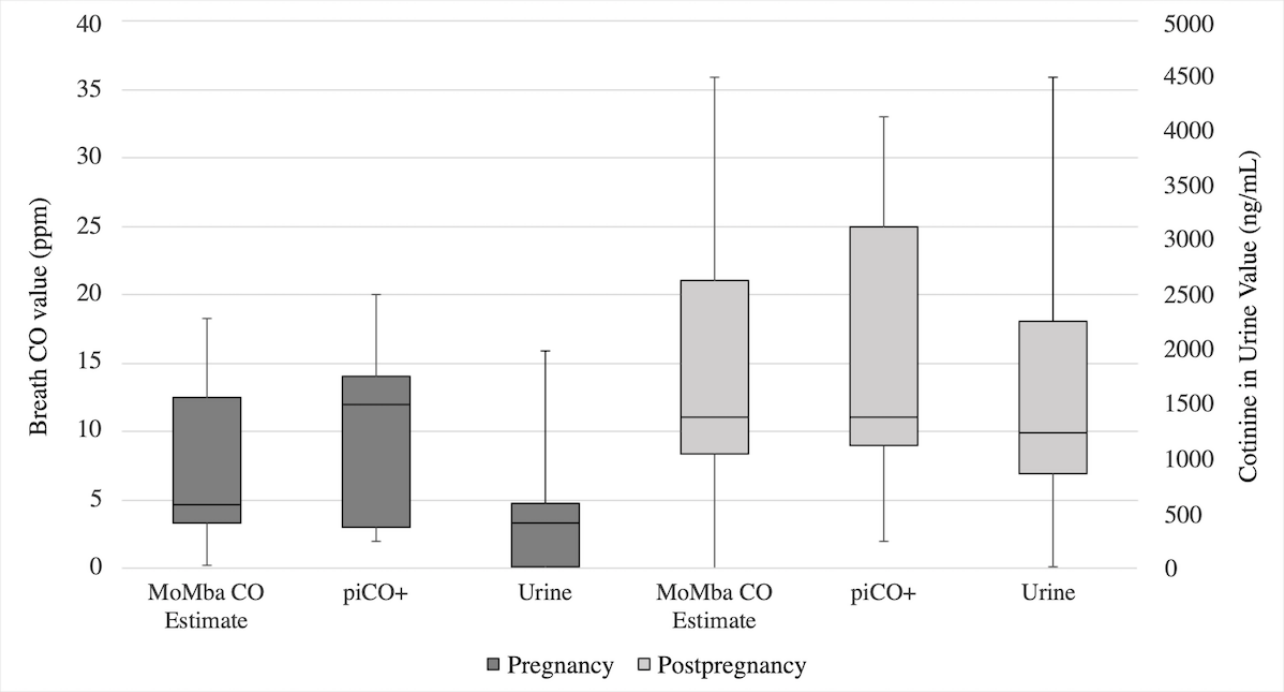
Figure 4. Breath carbon monoxide (CO) values according to time since last cigarette smoked. ppm: parts per million.



Variation of CO Values According to Pregnancy Status

Among the 4 participants who delivered a baby during the study (65 breath tests), the median MoMba CO estimate was 4.7 ppm (IQR 3.3-12.5) during pregnancy and 11.1 ppm (IQR 8.4-21.0) after pregnancy (see Figure 5). The median CO estimate from piCO$^+$ was 12 ppm (IQR 3-14) during pregnancy and 11 ppm (IQR 9-25) after pregnancy. There were no differences between values during pregnancy compared with values after pregnancy for all three smoking indicators; this remained after controlling for number of cigarettes smoked in the past 24 hours and time of last cigarette smoked.

Figure 5. Breath carbon monoxide (CO) and cotinine in urine values according to pregnancy status (n=4). ppm: parts per million.



Out-of-Office Performance and User Experience

Of the 102 remote breath tests successfully sent to participants, 51 (50.0%) tests were completed. Participants received, on average, 10 remote breath test notifications (SD 3.6) and completed 5 breath tests on average (SD 3.5). Almost all remote breath tests (47/51, 92%) indicated smoking. Remote breath tests that were not completed were preceded by a breath test indicating smoking 88% (45/51) of the time; average time from a missed remote test to the preceding completed breath test was 9.8 days (SD 7.2). Notifications were not successfully delivered for 7 out of 51 (14%) remote tests; in 2 of these instances (29%), the participant received a replacement breath test in the office, and in 2 instances (29%) a replacement breath test was sent remotely to the participant. Out of 51 remote breath tests, 4 (8%) results were challenged by a participant; research staff examined the data to determine if there was an inconsistency in breath sample values and pictures. If results were indicative of a false positive, the participant was sent a new breath test.

All 51 completed remote breath tests had pictures where staff could verify the identity of the participant completing the test; only one set of pictures for 1 remote breath test out of 51 (2%) did not show the participant's mouth. All participants used the front-facing camera feature correctly. Audio was available for all remote breath tests except for 1 out of 51 (2%), in which the file was corrupted. About 50% of audio recordings also included background noise, making it difficult to detect the sounds of breath; these background sounds included TV, radio, fans, babies, and other people talking.

At the 3-month follow-up, 63% (5/8) of participants indicated that the MoMba Live Long app and the Sensordrone were "extremely easy to use," and 38% (3/8) indicated that it was "easy to use."

Discussion

Principal Findings

In this study, we tested and validated the MoMba Live Long system's breath CO meter as a smartphone-based system that can determine smoking status during and after pregnancy. Collected breath CO values with the MoMba Live Long system compared well to the commercial piCO$^+$ monitor. Women in our study found the system easy to use. Regarding the novelty of our system, in comparison with other remote smoking measurement instruments, including the piCO$^+$ and the new iCO$^+$ Smokerlyzer (Bedfont Scientific), the MoMba Live Long CO meter has several advantages: (1) it interfaces wirelessly with a smartphone app, (2) it enables user verification, through picture and audio capturing for smoke test validity on the go, and (3) it can detect smoking status during and after pregnancy. Further investigation is required to determine what additional system features are needed to make home CM interventions feasible beyond ensuring the accuracy of the measurement system.

System Performance Metrics in the CM Context

When delivering financial incentives through remote verification of smoking status, specificity—the proportion of tests for which there is a true-negative breath CO test—is the measurement that should be prioritized to positively reinforce participants who have abstained from smoking. The MoMba Live Long system achieved high levels of specificity; false positives in the context of a CM intervention can decrease motivation to quit and increase frustration for participants.

Breath CO as a Time-Dependent Marker for Smoking

While CO concentration and cotinine can both measure the presence or absence of smoking, the difference in the half-lives of these analytes do not make them directly comparable. CO

has a shorter half-life (3-6 hours) in comparison to cotinine (17 hours in nonpregnant women and 9 hours in pregnant women [28]). This difference helps explain the finding of an observed, moderate, linear relationship between both the MoMba CO estimate and the piCO$^+$ measure with urine cotinine tests. CO concentration in the breath can be a better measure of cigarette consumption within a shorter period of time than cotinine, as shown in previous work [29], making it suitable for recent and immediate smoking assessments [12]. Furthermore, our analysis found differences in breath CO levels within the 5 hours of the last cigarette smoked and after the 5-hour mark, suggesting that future remote CM interventions should consider sampling more than one time a day.

The Need for Context-Aware CO Cutoffs

Prior work has suggested a variety of optimal cutoff breath CO levels to determine smoking status [12,24,30-32]. While we did not observe differences in CO values during pregnancy and the postpartum period, possibly due to a small sample size, the literature suggests that using different cutoff values during pregnancy and the postpartum period may be the best alternative, although there is no consensus on a set cutoff value. While one study suggests a CO cutoff of 2-3 ppm, other studies recommend a 4-ppm breath CO cutoff to identify pregnant smokers [24]. Studies reporting these cutoff values were based on self-report of smoking, which is not always a reliable measure [33-35]. The reported variability in cutoff values is endogenous to exhaled CO as a biomarker since it depends on the given environment in which the measurement is taken [36], an individual's physiology [11], the breath sampling procedure, and the breath CO measurement instrument [37]. Further validation on optimal breath CO cutoff levels is needed as well as systems that are flexible to adapt to cutoff changes through the perinatal period.

Limitations and Future Work

This study does not report on the success of financial incentives to prevent smoking relapse but rather on the validity of our measuring instrument and the participant's experience in using the MoMba Live Long system in person and remotely after a period of 3 months. Future publications will report on a randomized controlled trial on the prevention of smoking relapse using remotely delivered financial incentives. While participants reported that the system was easy to use, there are many unknowns regarding why participants missed remote breath tests. Additional research should investigate reasons for noncompletion, such as difficulty accessing their phone or the required sensor, recent smoking, the timing in which notifications were received, and the social context in which participants are asked to take a test. Prior work has reported some of these factors as negatively impacting adherence to remote dietary interventions [38].

Another limitation of the CO meter is that many other environmental factors can affect the readings of a remote CO monitor, such as air pollution, secondhand smoke, or the use of tetrahydrocannabinol (THC) or cannabis. Expired CO air from THC users has shown to double CO concentration levels [39]. Our sensor is susceptible to increased CO readings based on these external factors. Studies looking into enabling remote smoking detection should consider other substance use screening, as well as assessing a participant's environment to be aware of possible confounding factors.

Conclusions

The MoMba Live Long system is one of the first portable breath CO monitoring systems delivering remote CM smoking cessation interventions for pregnant women. The results from this study suggest that CO estimates derived from a smartphone-based breath CO meter are reliable and valid, but further testing in remote and diverse settings is needed to fully understand what environmental and usability barriers may impact the process of taking remote breath tests. Overall, the MoMba Live Long system is a feasible and acceptable approach to help practitioners and researchers increase access and delivery of CM smoking cessation interventions remotely to diverse populations.

Acknowledgments

This work was supported by the National Institutes of Health (NIH) through a grant from the National Cancer Institute (R01CA195654) with MS as principal investigator. FS is funded by an NIH award (K01 MH104739). The iOS app used for this study was developed by PreviewLabs. The authors gratefully acknowledge Mark Wagner, founder of Sensorcon and inventor of the Sensordrone; Michael Perlmutter and Brandon Hudik for the custom 3D-chassis design; and the members of the New Haven Moms Partnership for their help with participant recruitment and data collection.

Conflicts of Interest

None declared.

References

1. Higgins ST, Chilcoat HD. Women and smoking: An interdisciplinary examination of socioeconomic influences. Drug Alcohol Depend 2009 Oct 01;104 Suppl 1:S1-S5 [FREE Full text] [doi: 10.1016/j.drugalcdep.2009.06.006] [Medline: 19586726]
2. DiClemente C, Dolan-Mullen P, Windsor R. The process of pregnancy smoking cessation: Implications for interventions. Tob Control 2000;9 Suppl 3:III16-III21. [doi: 10.1136/tc.9.suppl_3.iii16] [Medline: 10982900]

3. Fingerhut LA, Kleinman JC, Kendrick JS. Smoking before, during, and after pregnancy. Am J Public Health 1990 May;80(5):541-544. [doi: 10.2105/ajph.80.5.541] [Medline: 2327529]
4. Clinical Practice Guideline Panel, Liaisons, and Staff. Treating Tobacco Use and Dependence: 2008 Update. Clinical Practice Guideline. Rockville, MD: US Department of Health and Human Services. Public Health Service; 2008 May. URL: https://www.aafp.org/dam/AAFP/documents/patient_care/clinical_recommendations/TreatingTobaccoUseandDependence-2008Update.pdf [accessed 2020-10-24]
5. Lumley J, Chamberlain C, Dowswell T, Oliver S, Oakley L, Watson L. Interventions for promoting smoking cessation during pregnancy. Cochrane Database Syst Rev 2009 Jul 08(3):CD001055 [FREE Full text] [doi: 10.1002/14651858.CD001055.pub3] [Medline: 19588322]
6. Mobile fact sheet. Pew Research Center. Washington, DC: Pew Research Center; 2019 Jun 12. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2020-10-20]
7. Chen Y, Madan J, Welton N, Yahaya I, Aveyard P, Bauld L, et al. Effectiveness and cost-effectiveness of computer and other electronic aids for smoking cessation: A systematic review and network meta-analysis. Health Technol Assess 2012;16(38):1-205, iii [FREE Full text] [doi: 10.3310/hta16380] [Medline: 23046909]
8. Carrasco-Hernandez L, Jódar-Sánchez F, Núñez-Benjumea F, Moreno Conde J, Mesa González M, Civit-Balcells A, et al. A mobile health solution complementing psychopharmacology-supported smoking cessation: Randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 27;8(4):e17530 [FREE Full text] [doi: 10.2196/17530] [Medline: 32338624]
9. Baskerville N, Struik L, Hammond D, Guindon G, Norman C, Whittaker R, et al. Effect of a mobile phone intervention on quitting smoking in a young adult population of smokers: Randomized controlled trial study protocol. JMIR Res Protoc 2015 Jan 19;4(1):e10 [FREE Full text] [doi: 10.2196/resprot.3823] [Medline: 25599695]
10. Meredith SE, Robinson A, Erb P, Spieler CA, Klugman N, Dutta P, et al. A mobile-phone-based breath carbon monoxide meter to detect cigarette smoking. Nicotine Tob Res 2014 Jun;16(6):766-773 [FREE Full text] [doi: 10.1093/ntr/ntt275] [Medline: 24470633]
11. Moscato U, Poscia A, Gargaruti R, Capelli G, Cavaliere F. Normal values of exhaled carbon monoxide in healthy subjects: Comparison between two methods of assessment. BMC Pulm Med 2014 Dec 16;14:204 [FREE Full text] [doi: 10.1186/1471-2466-14-204] [Medline: 25515007]
12. Javors MA, Hatch JP, Lamb RJ. Cut-off levels for breath carbon monoxide as a marker for cigarette smoking. Addiction 2005 Feb;100(2):159-167. [doi: 10.1111/j.1360-0443.2004.00957.x] [Medline: 15679745]
13. Stoops WW, Dallery J, Fields NM, Nuzzo PA, Schoenberg NE, Martin CA, et al. An internet-based abstinence reinforcement smoking cessation intervention in rural smokers. Drug Alcohol Depend 2009 Nov 01;105(1-2):56-62 [FREE Full text] [doi: 10.1016/j.drugalcdep.2009.06.010] [Medline: 19615830]
14. Adkins J, Dutta P. Monoxalyze: Verifying smoking cessation with a keychain-sized carbon monoxide breathalyzer. In: Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems (SenSys '16). 2016 Presented at: 14th ACM Conference on Embedded Network Sensor Systems (SenSys '16); November 14-16, 2016; Stanford, CA p. 190-201. [doi: 10.1145/2994551.2994571]
15. You C, Lin Y, Chuang Y, Lee Y, Hsu P, Lin S, et al. SoberMotion: Leveraging the force of probation officers to reduce the risk of DUI recidivism. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (to be presented at UbiComp 2018). 2018 Sep Presented at: ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (to be presented at UbiComp 2018); October 8-12, 2018; Singapore p. 1-34. [doi: 10.1145/3264956]
16. Hsu P, Lin Y, Huang J, Chang C, Lin S, Lee Y, et al. A mobile support system to assist DUI offenders on probation in reducing DUI relapse. In: Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2017 ACM International Symposium on Wearable Computers (UbiComp '17). 2017 Sep Presented at: 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2017 ACM International Symposium on Wearable Computers (UbiComp '17); September 11-15, 2017; Maui, HI p. 77-80 URL: https://dl.acm.org/doi/pdf/10.1145/3123024.3123154 [doi: 10.1145/3123024.3123154]
17. Dáu ALBT, Callinan LS, Mayes LC, Smith MV. Postpartum depressive symptoms and maternal sensitivity: An exploration of possible social media-based measures. Arch Womens Ment Health 2017 Feb;20(1):221-224. [doi: 10.1007/s00737-016-0650-4] [Medline: 27416930]
18. Wagner M. Sensordrone: A practical, tricorder like platform for consumers and mobile device developers. Tech Briefs. 2012 Jun 25. URL: https://contest.techbriefs.com/2012/entries/consumer-products/2649 [accessed 2020-10-25]
19. Valencia S, Smith MV, Atyabi A, Shic F. Mobile ascertainment of smoking status through breath: A machine learning approach. 2016 Presented at: 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON); October 20-22, 2016; New York, NY p. 1-7. [doi: 10.1109/uemcon.2016.7777917]
20. Higgins ST, Washio Y, Heil SH, Solomon LJ, Gaalema DE, Higgins TM, et al. Financial incentives for smoking cessation among pregnant and newly postpartum women. Prev Med 2012 Nov;55 Suppl:S33-S40 [FREE Full text] [doi: 10.1016/j.ypmed.2011.12.016] [Medline: 22227223]
21. Low ECT, Ong MCC, Tan M. Breath carbon monoxide as an indication of smoking habit in the military setting. Singapore Med J 2004 Dec;45(12):578-582 [FREE Full text] [Medline: 15568119]

22. Fagerström K. Determinants of tobacco use and renaming the FTND to the Fagerstrom Test for Cigarette Dependence. Nicotine Tob Res 2012 Jan;14(1):75-78. [doi: 10.1093/ntr/ntr137] [Medline: 22025545]
23. Middleton ET, Morice AH. Breath carbon monoxide as an indication of smoking habit. Chest 2000 Mar;117(3):758-763. [doi: 10.1378/chest.117.3.758] [Medline: 10713003]
24. Bailey BA. Using expired air carbon monoxide to determine smoking status during pregnancy: Preliminary identification of an appropriately sensitive and specific cut-point. Addict Behav 2013 Oct;38(10):2547-2550. [doi: 10.1016/j.addbeh.2013.05.011] [Medline: 23793041]
25. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J, et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011 Mar 17;12:77 [FREE Full text] [doi: 10.1186/1471-2105-12-77] [Medline: 21414208]
26. McCarthy WF, Guo N. The estimation of sensitivity and specificity of clustered binary data. In: Proceedings of SAS Users Group International 31 (SUGI 31). 2006 Presented at: SAS Users Group International 31 (SUGI 31); March 26-29, 2006; San Francisco, CA p. 1-10 URL: https://support.sas.com/resources/papers/proceedings/proceedings/sugi31/206-31.pdf
27. Irimata K, Wakim P, Li X. Estimation of correlation coefficient in data with repeated measures. In: Proceedings of the SAS Global Forum 2018. 2018 Presented at: SAS Global Forum 2018; April 8-11, 2018; Denver, CO URL: https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/2424-2018.pdf
28. Dempsey D, Jacob P, Benowitz NL. Accelerated metabolism of nicotine and cotinine in pregnant smokers. J Pharmacol Exp Ther 2002 May;301(2):594-598. [doi: 10.1124/jpet.301.2.594] [Medline: 11961061]
29. Secker-Walker RH, Vacek PM, Flynn BS, Mead PB. Exhaled carbon monoxide and urinary cotinine as measures of smoking in pregnancy. Addict Behav 1997;22(5):671-684. [doi: 10.1016/s0306-4603(97)00013-0] [Medline: 9347069]
30. Usmani ZC, Craig P, Shipton D, Tappin D. Comparison of CO breath testing and women's self-reporting of smoking behaviour for identifying smoking during pregnancy. Subst Abuse Treat Prev Policy 2008 Feb 17;3(1):4 [FREE Full text] [doi: 10.1186/1747-597X-3-4]
31. Hertzberg JS, Carpenter VL, Kirby AC, Calhoun PS, Moore SD, Dennis MF, et al. Mobile contingency management as an adjunctive smoking cessation treatment for smokers with posttraumatic stress disorder. Nicotine Tob Res 2013 Nov;15(11):1934-1938 [FREE Full text] [doi: 10.1093/ntr/ntt060] [Medline: 23645606]
32. Benowitz NL, Bernert JT, Foulds J, Hecht SS, Jacob P, Jarvis MJ, et al. Biochemical verification of tobacco use and abstinence: 2019 update. Nicotine Tob Res 2020 Jun 12;22(7):1086-1097. [doi: 10.1093/ntr/ntz132] [Medline: 31570931]
33. Shipton D, Tappin DM, Vadiveloo T, Crossley JA, Aitken DA, Chalmers J. Reliability of self reported smoking status by pregnant women for estimating smoking prevalence: A retrospective, cross sectional study. BMJ 2009 Oct 29;339:b4347 [FREE Full text] [doi: 10.1136/bmj.b4347] [Medline: 19875845]
34. Ford RP, Tappin DM, Schluter PJ, Wild CJ. Smoking during pregnancy: How reliable are maternal self reports in New Zealand? J Epidemiol Community Health 1997 Jun;51(3):246-251. [doi: 10.1136/jech.51.3.246] [Medline: 9229052]
35. Campbell E, Sanson-Fisher R, Walsh R. Smoking status in pregnant women assessment of self-report against carbon monoxide (CO). Addict Behav 2001;26(1):1-9. [doi: 10.1016/s0306-4603(00)00070-8] [Medline: 11196282]
36. Crowley TJ, Andrews A, Cheney J, Zerbe G, Petty TL. Carbon monoxide assessment of smoking in chronic obstructive pulmonary disease. Addict Behav 1989 Jan;14(5):493-502. [doi: 10.1016/0306-4603(89)90069-5]
37. Raiff BR, Faix C, Turturici M, Dallery J. Breath carbon monoxide output is affected by speed of emptying the lungs: Implications for laboratory and smoking cessation research. Nicotine Tob Res 2010 Aug;12(8):834-838 [FREE Full text] [doi: 10.1093/ntr/ntq090] [Medline: 20530193]
38. Robinson E, Higgs S, Daley AJ, Jolly K, Lycett D, Lewis A, et al. Development and feasibility testing of a smart phone based attentive eating intervention. BMC Public Health 2013 Jul 09;13:639 [FREE Full text] [doi: 10.1186/1471-2458-13-639] [Medline: 23837771]
39. Hecht E, Vogt TM. Marijuana smoking: Effect on expired air carbon monoxide levels. Int J Addict 1985 Feb;20(2):353-361. [doi: 10.3109/10826088509044917] [Medline: 4008128]

Abbreviations

AUC: area under the curve
CM: contingency management
CO: carbon monoxide
HELLP: hemolysis, elevated liver enzymes, low platelet count
NIH: National Institutes of Health
ppm: parts per million
THC: tetrahydrocannabinol

Edited by G Eysenbach; submitted 23.03.20; peer-reviewed by N Walker, J Chan, G Humphrey, G Wadley, K Bold, T Brandon; comments to author 12.06.20; revised version received 07.08.20; accepted 02.09.20; published 24.11.20.

Please cite as:
Valencia S, Callinan L, Shic F, Smith M
Evaluation of the MoMba Live Long Remote Smoking Detection System During and After Pregnancy: Development and Usability Study
JMIR Mhealth Uhealth 2020;8(11):e18809
URL: https://mhealth.jmir.org/2020/11/e18809
doi:10.2196/18809
PMID:33231550

Original Paper

Development and Acceptability of a Tablet-Based App to Support Men to Link to HIV Care: Mixed Methods Approach

Thulile Mathenjwa[1], MA; Oluwafemi Adeagbo[1,2,3], PhD; Thembelihle Zuma[1], PhD; Keabetswe Dikgale[1], BSc; Anya Zeitlin[4], MSc; Philippa Matthews[1], BMBS; Janet Seeley[1,5], PhD; Sally Wyke[6], PhD; Frank Tanser[1,7], PhD; Maryam Shahmanesh[1,8*], PhD; Ann Blandford[4*], PhD

[1]Africa Health Research Institute, KwaZulu Natal, Mtubatuba, South Africa

[2]Department of Sociology, University of Johannesburg, Johannesburg, South Africa

[3]Department of Health Promotion, Education and Behavior, University of South Carolina, Columbia, SC, United States

[4]UCL Interaction Centre, University College London, London, United Kingdom

[5]London School of Hygiene and Tropical Medicine, London, United Kingdom

[6]University of Glasgow, Glasgow, United Kingdom

[7]Lincoln International Institute for Rural Health, University of Lincoln, Lincoln, United Kingdom

[8]Institute for Global Health, University College London, London, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Thulile Mathenjwa, MA
Africa Health Research Institute
KwaZulu Natal
Somkhele, Africa Centre Building, Via R618 to Hlabisa
Mtubatuba, 4001
South Africa
Phone: 27 355507500
Email: thulile.mathenjwa@ahri.org

Abstract

Background: The poor engagement of men with HIV care is attributed to a number of factors: fear of stigma, masculine representations, concerns related to confidentiality, and the time commitment needed to visit public health clinics. Digital technologies are emerging as an approach to support the engagement of men with care.

Objective: This study aims to deliver a usable and engaging tablet-based app, called EPIC-HIV 2 (Empowering People through Informed Choices for HIV 2), to support men in making informed decisions about engaging with HIV care in rural KwaZulu Natal, South Africa.

Methods: We employed a mixed methods, iterative, and three-phased design that was guided by self-determination theory (SDT), a person-based approach, and human-computer interaction techniques. We reviewed related literature and conducted secondary analyses of existing data to identify barriers and facilitators to linkage to care and inform content development and design principles and used focus group discussions with members of the community advisory board and general community to evaluate a PowerPoint prototype of the app; used observations and guided questions with a convenience sample of potential users from the intervention community to iteratively test and refine a functioning interactive app; and conducted qualitative interviews and satisfaction surveys with actual users to evaluate acceptability.

Results: Phase 1 identified supply- and demand-side barriers to linkage to care. Specifically, clinics were feminized spaces unattractive to men with high social costs of attendance. Men did not feel vulnerable to HIV, preferred traditional medicine, and were afraid of the consequences of being HIV positive. Thus, the app needed to allow men to identify the long-term health benefits to themselves and their families of starting antiretroviral therapy early and remaining on it, and these benefits typically outweigh the social costs of attending and being seen at a clinic. SDT led to content design that emphasized long-term benefits but at the same time supported the need for autonomy, competence, and relatedness and informed decision making. Phase 2 indicated that we needed to use simpler text and more images to help users understand and navigate the app. Phase 3 indicated that the app was acceptable and likely to encourage men to link to care.

Conclusions: We found that iteratively developing the app with potential users using local narratives ensured that EPIC-HIV 2 is usable, engaging, and acceptable. Although the app encouraged men to link to HIV care, it was insufficient as a *stand-alone intervention* for men in our sample to exercise their full autonomy to link to HIV care without other factors such as it being convenient to initiate treatment, individual experiences of HIV, and support. Combining tailored digital interventions with other interventions to address a range of barriers to HIV care, especially supply-side barriers, should be considered in the future to close the present *linkage gap* in the HIV treatment cascade.

(JMIR Mhealth Uhealth 2020;8(11):e17549) doi:10.2196/17549

KEYWORDS

HIV; linkage to HIV care; digital technologies; men; mobile phone

Introduction

Poor Engagement of Men With HIV Care

Poor engagement of men with HIV care is well described [1-5]. This results in the persistence of unacceptably high HIV-related mortality among men and new HIV cases in young women [6]. In rural KwaZulu-Natal, South Africa, despite freely available antiretrovirals (ARVs) in the public sector, men are 25% more likely than women to die of HIV-related illnesses, and 70% of them had never sought care [7]. Fear of stigma, cultural constructions of masculinity that do not involve caring for health, concerns related to confidentiality, and time commitment needed to attend public health clinics have been cited as key reasons for men's reluctance to get tested and start HIV treatment [1,8]. Novel, scalable interventions that improve the uptake of HIV testing and care among men are needed to improve individual health and reduce HIV transmission risk [9,10].

The Potential of Digital Technologies in Closing Gaps in HIV Care

Digital technologies such as mobile phones are emerging as one of the approaches to close the gaps across the HIV care cascade [11,12], especially in light of the expanding access to smartphones in areas most affected by HIV/AIDS. A recent Pew internet survey found that 33% of respondents from sub-Saharan Africa owned a smartphone, and the number is expected to double by 2025 [13]. Digital technologies are particularly suitable as interventions to help prevent and manage HIV and AIDS, as they have the potential to access and impact hard-to-reach populations, including those who typically feel stigmatized within health care settings [14]. Daher et al [15] conducted a systematic review of digital innovations for HIV or sexually transmitted infections and found them to be effective in improving clinic attendance, adherence to antiretroviral therapy (ART), and turnaround time from testing to treatment. This suggests that digital interventions have considerable potential to support engagement in the HIV care continuum.

Challenges to the Development of Digital Interventions

There are many challenges to the development of digital interventions: most importantly, ensuring that the digital platform is usable and engaging for end users [16] and delivering the intended messages. Designing digital interventions for resource-constrained settings such as rural South Africa has added challenges, as there are varying levels of education, health, and digital literacy [17]. Furthermore, ensuring that the intervention supports individual decision making can be challenging. There are limited data on how to develop digital interventions for HIV care in resource-constrained settings that are usable and engaging and support informed individual decision making.

As part of a larger, cluster randomized controlled clinical trial to increase home-based HIV testing and linkage to care in men [18], we developed 2 interactive tablet-based apps (to integrate with the trial) called EPIC-HIV 1 and 2 (Empowering People through Informed choices for HIV 1 and 2). Both are guided by self-determination theory (SDT) to support informed decision making and make explicit and usually implicit processes of decision making. EPIC-HIV 1 is offered at the point of HIV test to help men make an informed choice about testing and if necessary linkage to care. EPIC-HIV 2 is offered to men who have tested HIV positive but have not linked to care within a specified time frame. It aims to encourage men to link to care and stay in care by supporting autonomy, feelings of relatedness, and competence in linking to care by using positive examples from other men living with HIV.

Theoretical Framework

The primary aim of EPIC-HIV 2 is to support informed decision making by making explicit decisions that are usually implicit and thus, encourage men to link to care, to initiate ART early, and to introduce the potential benefits of staying on ART. Therefore, it is aimed at shifting the motivation of link to care from unmotivated or a motivated to internally motivated because of the benefits of ART. SDT addresses factors that either facilitate or undermine motivation. According to SDT, human motivation is based on the satisfaction of 3 inherent psychological needs: (1) autonomy, (2) competence, and (3) relatedness. Autonomy requires an individual to make personal decisions and act in a way that corresponds to their identity, belief systems, and values [19,20]; competence refers to an individual's perceived ability to perform an act [21]; and relatedness refers to the ability of an individual to connect to others and feel cared for [22]. SDT proposes that events or conditions that enhance a person's sense of autonomy, competence, and relatedness support internal motivation, which, in turn, can facilitate the adoption of new behaviors to be internally motivated and sustained [20,21].

In the case of HIV, SDT has been used to understand ART adherence, treatment motivation [23], and overall health behavior among individuals living with HIV [22,24]. Applying evidence-based theories to the development process of

interventions helps to direct attention to design characteristics that may otherwise be ignored and indicate conditions under which interventions are more likely to be effective [25]. EPIC-HIV 2 development was guided by SDT to increase men's internal motivation to engage with HIV care. In this paper, we provide an overview of the approach we adopted for designing and developing EPIC-HIV 2. Using one component of the app as a detailed case study, we describe the design rationale and components of the app and report on its acceptability to users.

Methods

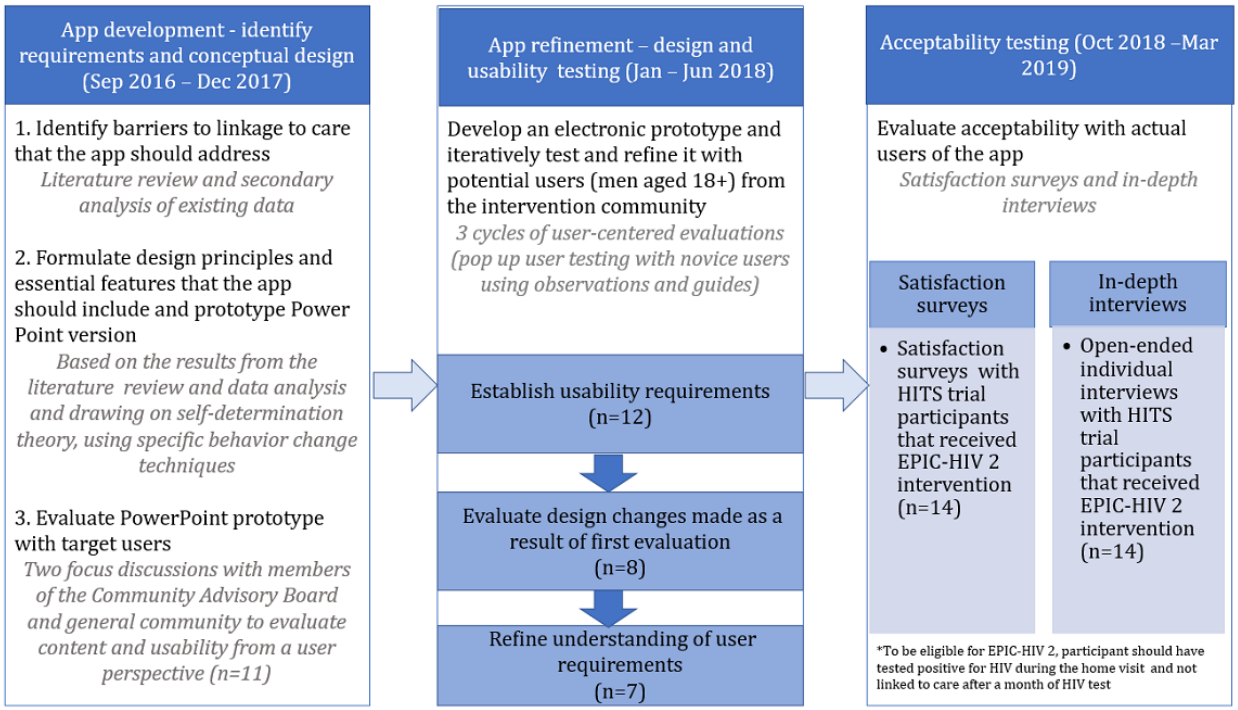
Study Context

As mentioned earlier, EPIC-HIV is being evaluated in a home-based intervention to test and start (HITS), a cluster randomized controlled trial designed to evaluate the effect of 2 interventions—EPIC-HIV (male-targeted HIV-specific decision support app) and microincentives on increasing home-based HIV testing and linkage to care among men, with the ultimate aim of reducing HIV-related mortality in men and hence HIV incidence in young women in rural South Africa [18]. The details of the study are described elsewhere [18] and is registered on the National Institute of Health trials (identifier NCT03757104). In summary, the trial uses the Africa Health Research Institute (AHRI) HIV surveillance platform, which visits households annually to conduct surveys using Android-based tablets and offer home-based HIV testing. EPIC-HIV 1 was administered at the point of HIV test offer during the fieldworker surveillance activity to encourage men to test for HIV. If a participant was diagnosed with HIV but did not present in the local Department of Health clinic within a month of the positive HIV test, a study fieldworker revisited the participant at home to offer EPIC-HIV 2. Both apps were installed on the field worker's tablet. During the home visit, the field worker handed over the tablet to the participant together with earphones to allow the participant to independently explore the app. EPIC-HIV 2 uses a mixture of audio, text, video, still photos, and graphics and has different pathways through 3 interactive modules. This paper focuses on the design and development of EPIC-HIV 2.

Study Design

This was a mixed methods study using a person-centered approach and human-computer interaction (HCI) techniques. The study was conducted in 3 phases. Phase 1 focused on user requirements and early prototyping. It included identifying barriers and facilitators to HIV linkage to care through literature review and secondary analysis of existing qualitative data from AHRI. We then developed a PowerPoint prototype of the app based on SDT and drawing on established behavior change techniques [26], which was discussed with members of the Community Advisory Board (CAB; special body representing members of the intervention community that act as a bridge between AHRI and the community, safeguarding the rights of the study participants) and general members of the community in focus group discussions (FGDs). Together, these methods were used to inform app content requirements and conceptual design. In phase 2, based on the findings from phase 1, we developed a functioning, interactive app. We iteratively tested and refined the app with potential users from the intervention community using observations and guided surveys. Finally, in phase 3, we evaluated the acceptability of the app with actual users (participants who received the EPIC-HIV intervention in the HITS trial). This process is illustrated in Figure 1.

Figure 1. Mixed methods approach showing the 3 phases of the study. EPIC-HIV: Empowering people through informed choices for HIV; HITS: home-based intervention to test and start.



Study Setting

This study was conducted in uMkhanyakude district in KwaZulu-Natal, South Africa. The area is predominantly rural and poor, with high levels of unemployment [27] and HIV prevalence [28]. Previous research has shown poor linkage and low ART uptake among men in this area [29,30] mainly because of fear of stigma [1,5].

Recruitment and Sample

All participants were recruited from the AHRI HIV surveillance community. Each phase employed different recruitment and sampling methods; these are described in detail under each phase.

Ethical Approval

Ethical approval for the study was granted by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal (BREC 398/16).

Below, we provide an overview of each phase, outlining its main aim, study procedures, and participants.

Phase 1: Requirements and Conceptual Design (App Development, September 2016 to December 2017)

Aim

The aim is to (1) gather the app content requirements and formulate design principles and features that the app should include by identifying modifiable barriers to linkage to care, (2) develop an initial prototype, and (3) evaluate it with potential users for overall impressions, appropriateness, and relevance to our setting.

Study Procedures

First, we drew on systematic reviews conducted on barriers and facilitators to testing and linkage to care and searched for additional studies in South Africa or elsewhere in sub-Saharan Africa (Multimedia Appendix 1 [1,31-43]). Second, we conducted a thematic secondary analysis from an AHRI study that explored community experiences and perceptions of tuberculosis and HIV transmission, treatment, and prevention [44] (Multimedia Appendix 1). The thematic coding frame also identified supply- and demand-side barriers to linkage to care, focusing on men's perspectives. The findings from the literature review and secondary analysis were summarized as barriers to linkage, and those that were potentially modifiable were identified (Multimedia Appendix 1). Third, we identified specific behavior change techniques from the Behavior Change Taxonomy [26] that could support feelings of autonomy, relatedness, and competence and thus were likely to shift motivations for linkage to care, support informed decision making, and make explicit processes that are usually implicit (such as avoiding the clinic). Finally, we developed a PowerPoint prototype of the app and evaluated it with focus groups using an FGD guide (semistructured questions, assessing the content and design and overall app impressions). Informed consent was obtained from all participants before beginning each focus group. FGDs were conducted in the AHRI offices in isiZulu (native language), audio recorded, transcribed, translated, summarized by TZ, and discussed with the development team (MS, PM, SW, TM, and TZ). They were analyzed using thematic analysis [45] to draw out key themes for design.

Participants

To obtain diverse views, we invited a purposive sample of men aged 18 years and above from the CAB and general community. We sought members representing periurban and rural communities in the surveillance area. In general, we aimed for men of different ages and occupying various positions in the community. For example, we recruited men from the taxi industry, church, liquor business, and formal and informal occupations, those who were unemployed, and those in the traditional health sector. The AHRI public engagement (PE) department assisted with the recruitment of participants. Through PE, the research team explained the study to the CAB, and the members who were willing to participate and provide informed consent were phoned for further engagement and to arrange a suitable time for the FGD. Of the 10 CAB members, 7 were included in the study. To recruit community members, a research assistant in the study approached men meeting the study inclusion criteria and explained the study, and those who were willing to participate provided their telephone numbers to be contacted to arrange a time for the FGD. Of 11 men who were approached, 4 participated in the study. We conducted 2 FGDs, one for the CAB and one for the community members.

Phase 2: Iterative Design and Usability Evaluations (App Refinement, January to June 2018)

Aim

The main aim of this phase is to iteratively design and test the app to refine it to be usable and engaging in our setting.

Study Procedures

We conducted 3 cycles of usability evaluations to assess participants' ability to successfully navigate the app, comprehend the educational content, and determine whether they found the app to be engaging and relevant. The first cycle was to establish usability requirements, the second cycle was used to evaluate the design changes made as a result of the first evaluation, and the last cycle was used to refine the understanding of the user requirements. User-centered evaluations were conducted with potential users using observations and guided surveys that consisted of a mix of open and closed questions to assess comprehension of the content and record usability issues. The questions were adjusted to suit the objectives of each cycle of evaluation with 2 questions that remained consistent to benchmark the design changes (Multimedia Appendix 2). Evaluations were conducted in isiZulu, and we recorded the audio and tablet screen during the session. Data files were translated and transcribed, summarized, and discussed with the HCI expert and technological partner after each evaluation cycle. On the basis of the findings, the app was tailored to ensure that participants understood the messages correctly and to improve the user experience.

Participants

We used pop-up user testing [46] to recruit a convenience sample of men aged 18 years and above from the intervention

community. We sought participants who represented a broad spectrum of technological and educational literacy, involving at least six participants in each cycle to get quick feedback on the app. The team drove (in an AHRI-marked vehicle) to different locations in the area and asked men (novice users) to participate in this study. Participants gave verbal consent and were asked to confirm if they were aged 18 years and above; they were not required to disclose their HIV status or any personal information.

Phase 3: Acceptability Testing (October 2018 to March 2019)

Aim

The aim of this phase is to evaluate the acceptability and perceived value of the resulting app with actual users (participants who received the EPIC-HIV intervention in the HITS trial).

Study Procedures

We conducted satisfaction surveys and individual in-depth interviews. An adapted Client Satisfaction Questionnaire (CSQ-8) [47] was used to assess satisfaction with the app (see Table 1 for the adapted questionnaire). The CSQ-8 has 8 items: quality of app, app met needs, kind of service received from the app, recommend app to a friend, amount of help received from the app, effectiveness of the app for dealing with the health problem, overall satisfaction, and willingness to use the app again. These items are assessed on a 4-point Likert scale of 1 to 4 with individually specified anchors, and 4 is consistently the positive assessment. For this study, we added 2 questions to assess the relevance of the app for this setting and user friendliness. The CSQ-8 has been used to evaluate technology-based intervention and has demonstrated high consistency [48]. The 10 items were administered on a REDCap [49] project by a fieldworker. Proportions were used to describe the acceptability. Statistical analyses were conducted using STATA software (version 15.1; StataCorp LLC). We complemented the satisfaction surveys with in-depth interviews with a purposive sample of participants who received EPIC-HIV 2. The interview questions explored why participants had not been linked to care after the HIV diagnosis and their views of EPIC-HIV 2 and the impact it had on them. Interviews were conducted at participants' homes in isiZulu and lasted approximately an hour. The audio files were transcribed, translated, and analyzed thematically with themes guided by SDT. NVivo was used to code and manage the transcripts. Separate informed consent was obtained for participating in the satisfaction surveys and in-depth interviews.

Table 1. Empowering People through Informed Choices for HIV 2 satisfaction surveys.

Survey item	Response, n (%)
How would you rate the quality of the EPIC-HIV[a] app you just listened to?	
Excellent	13 (93)
Good	1 (7)
Fair	0 (0)
Poor	0 (0)
To what extent has the app helped in meeting your needs with regards to information about HIV treatment?	
All	12 (86)
Most	0 (0)
Only a few	1 (7)
None	1 (7)
Did you get the kind of HIV management information you wanted or expected?	
Yes, definitely	14 (100)
Yes, generally	0 (0)
No, not really	0 (0)
No, definitely	0 (0)
Did the information you just listened to appeal to your conscience to go for HIV treatment and take ARVs[b] if you have not done so?	
Yes, definitely	13 (93)
Yes, generally	1 (7)
No, not really	0 (0)
No, definitely	0 (0)
Do you feel empowered to make informed choices regarding your health with the information from the app?	
Yes, definitely	12 (86)
Yes, I think so	2 (14)
No, I don't think so	0 (0)
No, definitely not	0 (0)
If a friend were in need of HIV treatment and management information, would you recommend the app to him or her?	
Yes, definitely	14 (100)
Yes, I think so	0 (0)
No, I don't think so	0 (0)
No, definitely not	0 (0)
How satisfied are you with the amount of HIV management information you received from the app?	
Very satisfied	12 (86)
Mostly satisfied	2 (14)
Mostly dissatisfied	0 (0)
Quite dissatisfied	0 (0)
How would you rate the simplicity and user-friendliness of the app?	
Excellent	11 (79)
Good	2 (14)
Fair	1 (7)
Poor	0 (0)
In general, how satisfied are you with the app?	

Survey item	Response, n (%)
Very satisfied	14 (100)
Mostly satisfied	0 (0)
Mostly dissatisfied	0 (0)
Quite dissatisfied	0 (0)
If you were to seek information about HIV management again, would you consider EPIC-HIV app?	
Yes, definitely	14 (100)
Yes, I think so	0 (0)
No, I don't think so	0 (0)
No, definitely not	0 (0)

[a]EPIC-HIV: Empowering People through Informed Choices for HIV 2.

[b]ARV: antiretroviral.

Participants

All men aged 15 years and above who received EPIC-HIV 2 between October 2018 and January 2019, as part of the intervention in the HITS trial, were asked to consent and complete the satisfaction surveys. For the individual in-depth interviews, we purposively selected 14 men from the 28 who had received EPIC-HIV 2 between April 2018 and January 2019 (which meant that they had tested positive for HIV but did not link to care within a month).

Results

Each phase is presented separately with a particular focus on the iterative development. Textbox 1 shows the key takeaway findings from the 3 phases.

Textbox 1. Key takeaway findings for the different phases.

Phase 1

- Identified key learning points to incorporate in the content to persuade men to attend clinic and start antiretroviral therapy
- Long-term health benefits of starting antiretroviral therapy early and remaining on it

Phase 2

- Men in our setting did not understand the seesaw metaphor
- The click worked better than swipe
- Two options were better than 3 options

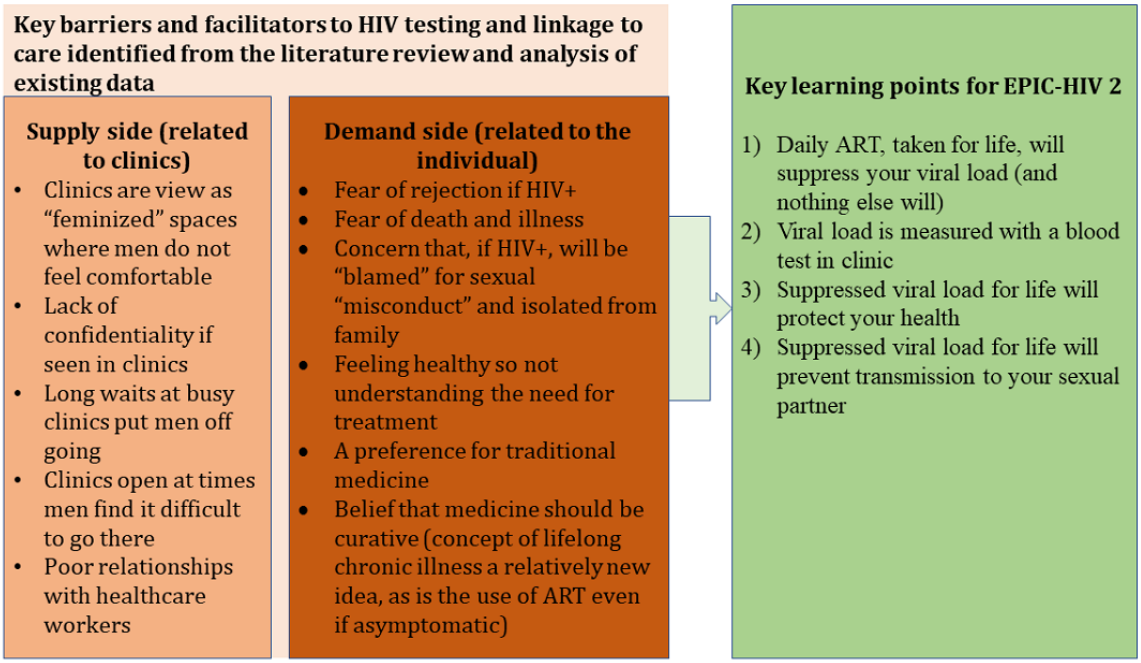
Phase 3

- App was usable and acceptable in our setting
- It encouraged men to link to HIV care

Phase 1: App Development

In the literature review and analyses of the earlier study, we identified barriers and facilitators to HIV testing and linkage to care among men, particularly relevant to our setting (Figure 2). Through this process, we identified 4 key learning points (shown in Figure 2) to be incorporated into the content of EPIC-HIV 2 with the aim of supporting informed decision making by making explicit decisions that are usually implicit and demonstrate that attending a clinic to start ART can be managed and that the long-term benefits of starting ART early and remaining on ART outweigh the costs. Drawing on SDT, we developed the app content and design to map to the psychological requirements for autonomy, relatedness, and competence (shown in Multimedia Appendix 3). To increase the likely personal identification with the messages and potential feelings of relatedness, the content was provided in the form of personal testimonies or experiences from local men. This was informed by evidence that suggests experiential information helps users make sense of what various outcomes might be like (in imagined futures) and increases their awareness of personal health risk and the likelihood of a response to the intervention [50].

Figure 2. The identified key barriers and facilitators to HIV testing and linkage to care and the key learning points for EPIC-HIV 2. ART: antiretroviral; EPIC-HIV: Empowering people through informed choices for HIV; HITS: home-based intervention to test and start.



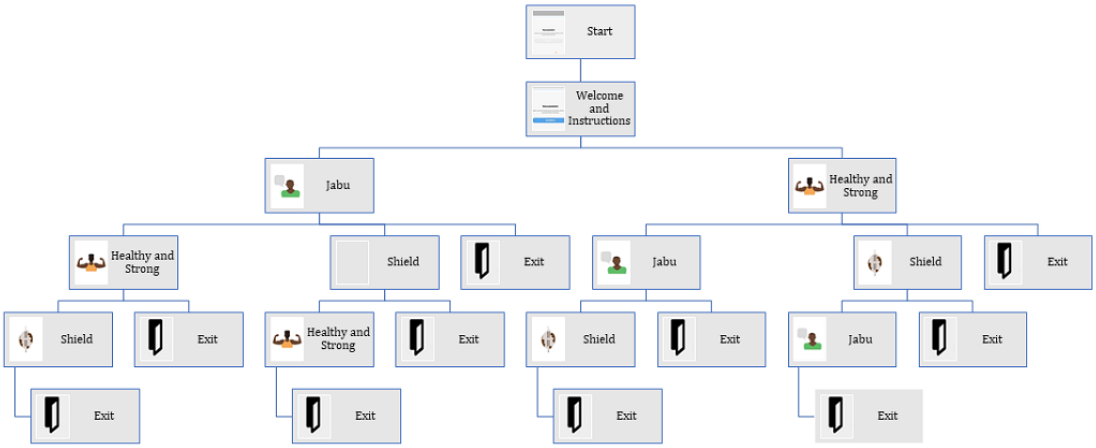
A total of 11 men aged 34 to 66 years participated in 2 focus groups to evaluate the first (paper) prototype. They provided feedback on content and design (organization of the app, appropriateness and appeal of language and images, acceptability and relevance, and overall app impressions). On the basis of the feedback from the FGDs, the team modified the initial conceptual design of the app and created a revised plan for the app features and content.

The final realized product (EPIC-HIV 2) has 3 main modules (sections), and to further support the psychological requirements for autonomy, the user has the option to engage with 1 or all 3 modules (Figures 3 and 4; Multimedia Appendix 3). All 3 modules cover the 4 key learning points (Figure 2) but present the information in different ways to support different learning styles, social identities, and barriers to engaging with care: (1) Jabu's man-to-man advice—a taxi driver talking about how he views HIV as an uninvited guest and gives his personal testimony on how he controls HIV; (2) healthy and strong with HIV—4 male characters who give their experiential information on how they managed going to the clinic, how they disclosed their HIV status, and how they started ART, and stayed on ART; and (3) the facts are your shield—gives a broader picture of ART and how it fights HIV in the body by explaining the commonly used medical terminology in HIV care. There is a start button, and a male voice over that introduces EPIC-HIV 2 and gives instructions for navigating the app. As mentioned earlier, the app uses a mixture of audio, text, video, still photos, and graphics, and there is a continue button that the user needs to click to move to the next page (Multimedia Appendix 4 provides an example of the navigation in the Jabu module).

The final conceptual design (realized in the product) starts with an introduction and instructions on how to use the app, and users are then able to choose between "Jabu's man to man advice" or "healthy and strong with HIV" modules. The "shield" module was not offered initially because it has more focused material and is expected to appeal to fewer men (and to make the number of choices manageable). After users complete the initial module, they are returned to the menu and are able to choose one of the remaining modules or to end. Figure 4 shows all the possible pathways through the app. Each module was developed and tested individually, as described in phase 2, and once finalized, all the different modules were integrated into a single app that was evaluated for functionality and comprehension.

Figure 3. PowerPoint prototypes showing examples of the content from each module (presented in English, although all content was delivered in isiZulu in the final app).

Figure 4. Map of the pathways through the 3 main modules of empowering people through informed choices for HIV 2.



Phase 2: Iterative Design and Usability Testing (January to June 2018)

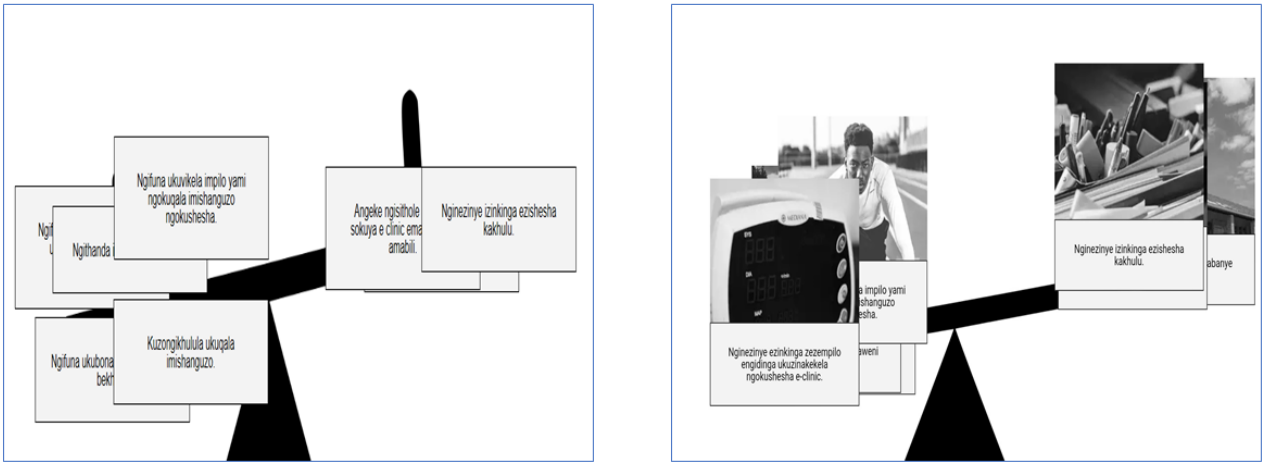
Decision Support Tool

The main outcome of phase 1 was an in-depth understanding of user needs relevant to EPIC-HIV 2. We subsequently refined the prototypes using a standard app development toolkit to deliver testable interactive demonstrator prototypes on an Android tablet computer. In this paper, we illustrate the approach to iterative design and usability testing by focusing on the development of the decision support tool (a component within *Jabu*).

The decision support tool aimed to help men weigh up and reflect on the costs and benefits of attending a clinic. The user was presented with different statements that either support or delay going to the clinic. To facilitate engagement, we wanted to make this section interactive. Initially, we applied a seesaw metaphor for users to attempt to render usually implicit decisions as explicit and interpret whether they were ready to attend a clinic within 2 weeks. This was done by adding the arguments to one side or the other of the seesaw to illustrate how the arguments stacked up in decision making. For example, if the user agrees that they want to see their children and grandchildren grow up, that would be an argument that weighs in favor of engaging with HIV care, whereas if they agree that they do not wish to be seen at a clinic by one of their neighbors, that would be an argument that weighs against engaging with care. In making design decisions about the clinic decision support tool, we prototyped and tested 2 versions: drag-and-drop with illustrations and a Likert scale without illustrations, both employing the seesaw metaphor. These are illustrated in Figure 5 (loose translation: “ngifuna ukubona abazukulu bami”—I want to see my grandchildren; “qiniso”—true; “amanga”—false; “ngiyavumelana”—I agree; “ngiphakathi nendawo”—neutral; “angivumelani”—I disagree) and Figure 6.

Figure 5. Example screenshots of drag-and-drop (statement with an image and illustration on how to drag) versus Likert scale (statement only).

Figure 6. Example of the seesaw screenshots with statements only (from Likert scale version) versus statements with images (from drag-and-drop version).



Evaluation 1

The main objective of the first evaluation was to refine the user requirements. The key objectives and results of this evaluation are shown in Textbox 2.

In total, 12 men participated in this evaluation, which was conducted as outlined earlier. In addition to the questions related to the key objectives outlined in Textbox 2, we also asked participants what the main messages were and whether there were any statements that should be added (ie, other considerations in deciding whether to engage in HIV care) and whether there were other ways the design could be improved. Overall, participants found it difficult to use the drag-and-drop option and did not understand the seesaw metaphor. Clicking was intuitive, and most participants preferred the images and text and the 2 options (agree or disagree). There was a consensus that the key message of this section was going to the clinic to start ART.

Textbox 2. Key objectives and results of evaluation 1.

To test which interaction style participants preferred and found easier to use between a drag-and-drop version (dragging each statement into a bucket labeled *agree* or *disagree*) or a Likert scale version (where people clicked in a box labeled *agree*, *neutral* or *disagree*")

- Participants found it difficult to use the drag-and-drop version, particularly more aged participants and participants with lower digital literacy levels
- A majority of participants needed to be shown how to drag-and-drop, despite the illustration at the beginning of the statements
- In contrast, clicking was intuitive; almost all the participants started clicking when they took the tablet

To determine whether the *neutral* option was valued by participants

- Some participants preferred drag-and-drop because it had 2 options (agree or disagree): "drag—I found it easy with the scale I was not completely sure of what was happening. With the drag option there was only two options which made it easy to select" [P6]
- Another participant reported that he found the neutral option confusing

To test whether participants preferred each option presented as text only or as text with an illustrative picture

- Most participants preferred text with an illustrative image to help them select the statements over the text only with many participants stating that images made it easier to understand the statements: "the image makes it clear what the statement says" [P7]
- Some found the images to facilitate engagement: "the images capture your attention" [P6]

To test whether participants were able to correctly interpret a seesaw metaphor, in which each response was stacked up on the end of a seesaw depending on whether it weighed toward or against attending a clinic (correct interpretation of the metaphor was that the end that was more heavily weighted and hence *down* was the preferred outcome)

- Participants did not understand the seesaw metaphor
- Many expressed confusion on how to interpret the results: "not sure how it should look like on the seesaw. When is it good, when it is up or down?" [P1]

To assess whether it was clear how to agree or disagree with a statement and whether it was clear how their selections related to the outcome of the seesaw

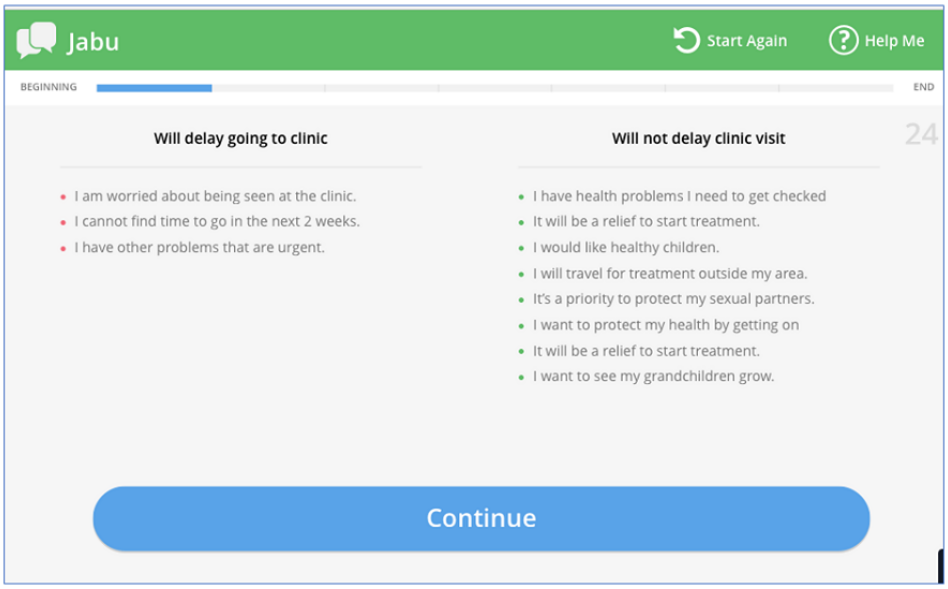
- It was not clear for some participants on how to agree or disagree with the statements; 2 participants (2/8) in the Likert scale option selected agree for all the statements

On the basis of the findings, the following design changes and decisions were made (as illustrated in Figures 7 and 8):

1. As the app was intended for one-off use, after seeing that clicking was intuitive and users struggled with the drag-and-drop, we decided to use clicking to select statements.
2. We limited the options to agree and disagree, as some participants did not find value in the neutral option.
3. To support users to make informed choices regarding attending a clinic for care, we used text and image as users expressed that images made it easier to understand the statements.
4. It was judged that the seesaw metaphor was not applicable for this exercise, as many participants struggled to understand what it meant if one side was more heavily weighted than the other. We removed the seesaw, summarized the statements, and explained what it means if they were skewed on one side. Furthermore, we labeled the right side as supporting going to the clinic in 2 weeks and left as delaying going to the clinic.
5. No further changes were judged necessary to address this question.

Figure 7. Example screenshot showing the design change with the 2 options, agree or disagree.

Figure 8. Example screenshot showing statements only without the seesaw and a clear label showing which side supports going to the clinic and which supports delaying (in English, but in isiZulu in the final app). This is a preliminary version, and the data were not final.



Evaluation 2

The main objective of the second evaluation is to validate the design changes made, particularly the summary screen shown in Figure 7. A total of 7 men participated in this evaluation cycle. We stopped them at the end of each section to ask guided questions: (1) Does the text only make the choices you selected clearer or do you think text and image would make it clearer? (2) What do you think is the key message of this section? (3) Would you prefer the agree and disagree buttons to be color coded?

Results of Evaluation 2

The design of the selection tool was received positively by all participants. All participants were selective on the statements they agreed or disagreed with as per their individual experiences, and they were able to interpret what it means if statements were weighted on one side (Figure 8). However, 2 participants mentioned that the statements were not clear or applicable to them. This could have been because the user was not affected by or living with HIV (the researcher did not ask the participant's HIV status).

Of 7 participants, 4 preferred the text only in the summary page, validating the first evaluation findings that images were only important at the selection point. In addition, 6 participants reiterated that the key message of this section was about the importance of going to the clinic to start ART, with 1 participant adding that one should go to the clinic early and not wait until they are sick. Moreover, 5 participants said they would prefer the agree or disagree buttons to be color coded, with 4 participants suggesting that agree should be green and disagree, red:

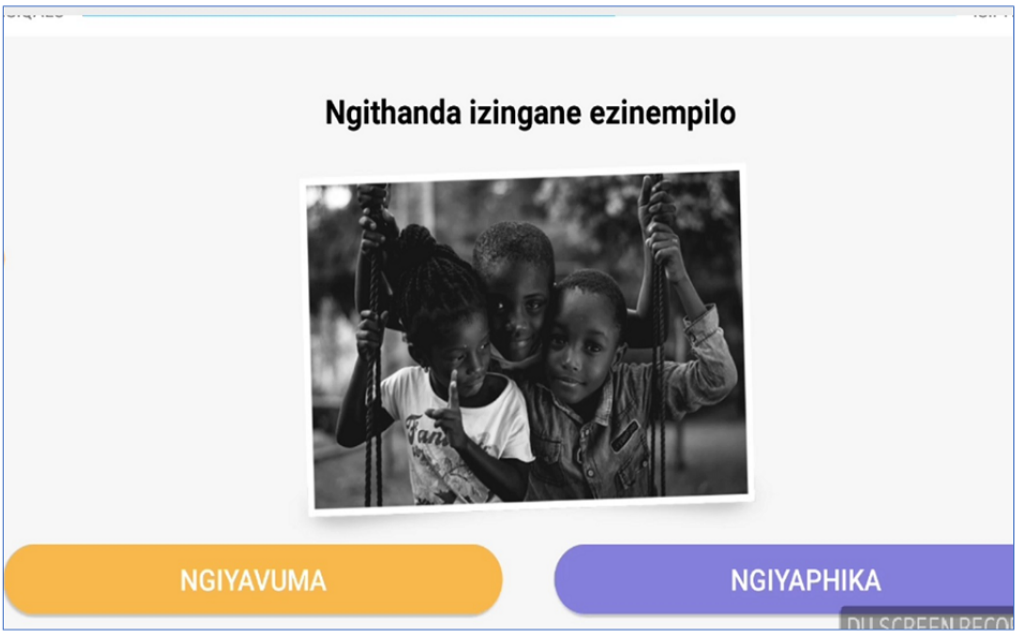
> *colours help, maybe green and red so that they can see that they have made the right or wrong choice.* [Participant 2]

One participant stated that different colors would be particularly useful for people who cannot read.

On the basis of the preference for color coding and the association of green with right and red with wrong, we decided

to use neutral colors that are not associated with right or wrong to help users make an individual choice as opposed to fearing being right or wrong (Figure 9).

Figure 9. Design changes showing the neutral colors for agree and disagree.



Evaluation 3

This evaluation was conducted as part of the full Jabu section testing, mainly to assess understanding of the content as well as to check whether users find the presentation of content interesting and to validate the changes made as a result of evaluation 2. A total of 8 men participated in this cycle.

This evaluation did not reveal any usability issues, indicating that the changes made to the designs were effective. Participants were paying more attention before deciding whether they agreed or disagreed, and they listened to the audio while looking at the screen. Overall, participants reported finding the decision support tool engaging and useful. For example, when asked which part of the app they thought would be most relevant to their friends or other men, 3 participants referred to the clinic decision support tool. In addition, 1 participant, when asked if he learned anything new as a result of using the app, mentioned that he learned how to use a tablet, suggesting that there was incidental value (not directly related to the primary aim) in introducing such an app to people.

Phase 3: Acceptability Testing (October 2018 to March 2019)

User Satisfaction Surveys

A total of 15 participants received EPIC-HIV 2 between October 2018 and January 2019, and 14 completed the user satisfaction surveys. Their mean age was 35.6 years (SD 13.38299), and a majority (8/14, 57%) had high school education, and half of them were unemployed. User satisfaction ratings were high (Table 1). Participants reported that they were generally satisfied with the app and found it to meet their needs for information on HIV treatment. All participants said that they received the kind of HIV management information that they wanted or expected and that they would definitely consider the app when seeking information about HIV management as well as recommending the app to friends in need of similar help. In addition, 86% (12/14) of participants were very satisfied with the amount of information that they received from the app. Overall, 79% (11/14) of the participants were very satisfied with the app by rating the simplicity and user friendliness of the app as excellent. The app has proven to empower the participants in managing their health; 86% (12/14) of participants definitely agreed.

In-Depth Interviews

A total of 14 men (aged 35-53 years) who had tested positive for HIV and had not linked to care within a month of receiving their test results received EPIC-HIV 2 between April 2018 and January 2019 agreed to be interviewed; 11 of them had linked to HIV care at the time of the interview. A majority of participants reported that the app offered relevant information about HIV management, boosted their confidence, and encouraged them to start and stay on treatment. However, this was not enough for all participants to link to HIV care. We report the experiences of men who used EPIC-HIV 2 using the SDT framework: autonomy, competence, and relatedness.

Complexities of the HIV Care Pathway: Examining Men's Autonomy to Link to HIV Treatment

Participants described the app as motivating them to start ART by providing them with information on how to manage HIV to live a healthy life:

> *Mmm... I can say I have listened to the App from the beginning up to the end. Yes, I have really seen a need...I have seen it as a very important thing. I think that has encouraged me because I have never been to the clinic previously, but I went to clinic after that.* [Participant 2, linked to care]

For some, this was a catalyst to link to care (even if they did not act on it immediately):

> *I was supposed to visit the clinic a long time ago, but I didn't because of work...I was thinking about that when I listened to EPIC [app]...you must be open as the EPIC was saying that you mustn't fear that people will laugh at you. I took a decision because...this is life and we are living within a community...I realised that in order to be assisted, I must go to AHRI nurses at the clinic.* [Participant 5, not linked to care]

However, for some participants, the app alone was insufficient to motivate them to go to the clinic for HIV treatment. Other factors reduced the logistic barriers to linkage that nudged them to link. For some participants, it was the availability of our research nurses at local clinics, with the:

> *referral letter... that was going to make things easy for me...* [Participant 4 linked to care]

For others, such as participant 1, it was experiences such as the death of his best friend that motivated him to link to care after using the app:

> *Er...I can just say that there is nothing new I have learnt (from EPIC)...I started treatment because I've lived with people who had been using it (ARVs) before I knew my health status. One of my best friends passed away after he defaulted from taking treatment. After I found out that I am HIV positive I was motivated, and I said I will try by all means to take treatment accordingly. So, I can say I have learnt a lot from him about the risk of delaying...* [Participant 1, linked to care]

This suggests that for men to exercise autonomy to link to care, in addition to the app, external factors such as clinic operation hours, support, and individual experiences of HIV need to be considered.

"I Am Not Giving Up on Life": Men's Competency and Decision to Manage HIV

Several of the participants reflected on how the positive narratives in the app encouraged them to take charge of their lives and manage HIV to improve their health outcomes and "avoid infecting other people with diseases" by describing "how to take the pills on time" and "comply with treatment."

Some participants also reported that app messaging lessened their fears (particularly fear of disclosure) and motivated them to adopt positive attitudes to advance their health. The quotation below illustrates how some participants felt about their fears:

> *...I had that feeling of fear to say, oh my Lord how could I do this thing since I am afraid as it will be revealed? Eh, I also gained advices because they start with allowing you to listen from the messages that say once you know your status, you can get help so it's up to you whether you do it or not. But I chose to be assisted because I am not giving up on life.* [Participant 4, linked to care]

Others also drew attention to the message of EPIC-HIV 2 that it gave them tips on how to discuss disclosure with their sexual partners:

> *Er...like disclosing your health status to your partner so that she can check for HIV as well for both of you to take treatment, yes, I think it's good not to hide from your partner:* [Participant 2, linked to care]

Overall, for many participants, the app supported their competence to link to care through the positive messaging that gave hope, alleviated their fears, and reinforced the need for them to manage their health effectively to live a healthy life while minimizing the risk of transmitting HIV.

Relatedness to Characters in the App: Men's Perceptions of HIV Treatment and Management

Participants reflected how the narratives of different characters in the app reverberated with their present realities and how the app stories shaped their understanding of the importance of HIV treatment and management for them to live a healthy life:

> *...They [app characters] have mentioned the fact that once you have defaulted from taking treatment you will be very sick…the characters I have seen on the app were fit. They were sharing their experiences as to how they are taking their ARVs...* [Participant 6, linked to care]
>
> *They [app characters] were sharing simple information, something that is said even on TV by Ministers that you need not to be bedridden until you go to the clinic.* [Participant 3, linked to care]

Furthermore, a participant who had previously defaulted treatment had this to say:

> *The part I have seen as most essential one for me it's where the characters were talking about the risk of defaulting from taking treatment...It motivated me to decide to be re-initiated on treatment...It is really necessary...because characters on the Epic were talking about things they had experienced.* [Participant 2, linked to care]

In sum, the empathetic nature of the personal testimonies of various characters in the app resonated with participants, and they felt cared for. Some of them described the relevance of the information shared through the app and were able to initiate treatment and took the decision to manage HIV to improve their health outcomes.

Discussion

Principal Findings

We found that it was possible to successfully apply a multiphased iterative development process to create a theoretically informed, interactive, tablet-based app to support men in making informed decisions about engaging with HIV care in a low-income, rural South African setting. Men in our setting who had been missing from the HIV care cascade found the app to be acceptable and reported that the stories in the app resonated with their realities and encouraged them to link to

care. Previous work has established that digital interventions are acceptable and effective in improving clinic attendance, ART adherence, and turnaround time from testing to treatment [11,14,15]. Our results underscore the value of using a person-based approach to integrate evidence-based content and design to ensure that the app is relatable and addresses the local perception of HIV care. Using HCI techniques ensured that the app was simple, usable, and engaging for end users with varying levels of education, health, and digital literacy. For example, user testing established that the *seesaw* metaphor was difficult for participants to understand, so it was removed between the first and the second iterations of app testing. In addition, supporting different learning styles and motivations through tailoring of content provided a greater level of engagement for participants. Previous studies have shown that interactive digital interventions that deliver tailored content that address the specific challenges for individual users can be highly engaging and likely to be understood [51]. Furthermore, using digital technologies could help people learn new skills (using a tablet) and facilitate a sense of accomplishment that can improve their competence and possibly improve engagement.

Drawing on behavior change theory, such as SDT, ensured that EPIC-HIV 2 supported individual decision making. There is an increasing need to apply evidence-based theories to enhance the development of digital interventions and improve their efficacy [15,25,52]. Our findings corroborate previous studies that have used the SDT framework (autonomy, competence, and relatedness) to understand ART adherence and treatment motivation and overall health behavior among men living with HIV [23,24]. However, the app was insufficient as a *stand-alone intervention* for men in our sample to exercise their full autonomy to link to HIV care without other factors such as it being convenient to initiate treatment, individual experiences of HIV, and support. Combining tailored digital interventions with other interventions to address a range of barriers to HIV care, especially supply-side barriers, should be considered in the future to close the present *linkage gap* in the HIV treatment cascade.

Study Strengths and Limitations

Using a mixed methods approach to understand usability and acceptability and the potential of the app to encourage men to link to care are strengths of the study. However, the generalizability of study findings outside our area may be limited because of the locally tailored design of the EPIC-HIV 2 app, the sampling method, sample size, and specific study sites. In addition, the possibility of social desirability bias cannot be excluded as some participants might have found it easier to say positive things about the app and not necessarily translate to the actual initiation of ART. Furthermore, we cannot establish what the outcome would have been if we had made different design decisions or applied a different theory of behavior change. The decision to develop a tablet-based app was fitting for the specific context of use, but more widespread use of such an app would probably be achieved by implementing it on mobile phones for independent use; this remains an area for future work.

Conclusions

Using a multidisciplinary approach and drawing on evidence-based theories to develop digital interventions can ensure that the resultant products are acceptable and engage in a wide range of users. Our work aims to pave the way for a greater focus on mixed methods and person-centered approaches to the development of digital interventions for HIV care.

Acknowledgments

The authors would like to acknowledge the study participants for their time. They wish to appreciate Velaphi Gumede, Sithembiso Luthuli, Siyabonga Nxumalo, and Jaco Dreyer for participant recruitment, data collection, and data management support. This study was supported by a Partnership Research Fund from the Engineering and Physical Sciences Research Council Interdisciplinary Research Collaboration Early-warning Sensing Systems for Infectious Diseases (i-sense) EP/K031953/1. The HITS trial is funded by the National Institutes for Health R01AI124389. The AHRI's Demographic Surveillance Information System and Population Intervention Program is funded by the Wellcome Trust (201433/Z/16/Z) and the South Africa Population Research Infrastructure Network (funded by the South African Department of Science and Technology and hosted by the South African Medical Research Council). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding bodies.

Authors' Contributions

FT, MS, SW, and JS are the investigators of the HITS trial and developed the EPIC-HIV intervention design and protocol in collaboration with AB and PM. PM, AB, AZ, TM, TZ, SW, MS, OA, and FT were involved in the development of EPIC-HIV. SW, TZ, MS, OA, and JS designed the process evaluation. KD provided statistical support. TM wrote the first draft of the manuscript with support from AB and OA and coordinated the study. All authors contributed to the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Literature review and secondary analysis.

[PDF File (Adobe PDF File), 243 KB - mhealth_v8i11e17549_app1.pdf]

Multimedia Appendix 2
User evaluation questions and guides.
[PDF File (Adobe PDF File), 108 KB - mhealth_v8i11e17549_app2.pdf]

Multimedia Appendix 3
Self-determination theory app content and design.
[PDF File (Adobe PDF File), 221 KB - mhealth_v8i11e17549_app3.pdf]

Multimedia Appendix 4
Example of navigation pathway in the Jabu module.
[PNG File , 286 KB - mhealth_v8i11e17549_app4.png]

References

1. Chikovore J, Gillespie N, McGrath N, Orne-Gliemann J, Zuma T, ANRS 12249 TasP Study Group. Men, masculinity, and engagement with treatment as prevention in KwaZulu-Natal, South Africa. AIDS Care 2016;28(Suppl 3):74-82 [FREE Full text] [doi: 10.1080/09540121.2016.1178953] [Medline: 27421054]
2. Govindasamy D, Ford N, Kranzer K. Risk factors, barriers and facilitators for linkage to antiretroviral therapy care: a systematic review. AIDS 2012 Oct 23;26(16):2059-2067. [doi: 10.1097/QAD.0b013e3283578b9b] [Medline: 22781227]
3. Kranzer K, Lewis JJ, Ford N, Zeinecker J, Orrell C, Lawn SD, et al. Treatment interruption in a primary care antiretroviral therapy program in South Africa: cohort analysis of trends and risk factors. J Acquir Immune Defic Syndr 2010 Nov;55(3):e17-e23 [FREE Full text] [doi: 10.1097/QAI.0b013e3181f275fd] [Medline: 20827216]
4. Tsai AC, Siedner MJ. The missing men: HIV treatment scale-up and life expectancy in sub-Saharan Africa. PLoS Med 2015 Nov;12(11):e1001906 [FREE Full text] [doi: 10.1371/journal.pmed.1001906] [Medline: 26599825]
5. Adeagbo O, Herbst C, Blandford A, McKendry R, Estcourt C, Seeley J, et al. Exploring people's candidacy for mobile health-supported HIV testing and care services in rural KwaZulu-Natal, South Africa: qualitative study. J Med Internet Res 2019 Nov 18;21(11):e15681 [FREE Full text] [doi: 10.2196/15681] [Medline: 31738174]
6. Cornell M, Schomaker M, Garone DB, Giddy J, Hoffmann CJ, Lessells R, International Epidemiologic Databases to Evaluate AIDS Southern Africa Collaboration. Gender differences in survival among adult patients starting antiretroviral therapy in South Africa: a multicentre cohort study. PLoS Med 2012;9(9):e1001304 [FREE Full text] [doi: 10.1371/journal.pmed.1001304] [Medline: 22973181]
7. Bor J, Rosen S, Chimbindi N, Haber N, Herbst K, Mutevedzi T, et al. Mass HIV treatment and sex disparities in life expectancy: demographic surveillance in rural South Africa. PLoS Med 2015 Nov;12(11):e1001905; discussion e1001905 [FREE Full text] [doi: 10.1371/journal.pmed.1001905] [Medline: 26599699]
8. Ahmed S, Autrey J, Katz IT, Fox MP, Rosen S, Onoya D, et al. Why do people living with HIV not initiate treatment? A systematic review of qualitative evidence from low- and middle-income countries. Soc Sci Med 2018 Sep;213:72-84 [FREE Full text] [doi: 10.1016/j.socscimed.2018.05.048] [Medline: 30059900]
9. Tanser F, Bärnighausen T, Grapsa E, Zaidi J, Newell M. High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. Science 2013 Feb 22;339(6122):966-971 [FREE Full text] [doi: 10.1126/science.1228160] [Medline: 23430656]
10. Rodger A, Cambiano V, Bruun T, Vernazza P, Collins S, Degen O, PARTNER Study Group. Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive partner taking suppressive antiretroviral therapy (PARTNER): final results of a multicentre, prospective, observational study. Lancet 2019 Jun 15;393(10189):2428-2438 [FREE Full text] [doi: 10.1016/S0140-6736(19)30418-0] [Medline: 31056293]
11. Simoni JM, Kutner BA, Horvath KJ. Opportunities and challenges of digital technology for HIV treatment and prevention. Curr HIV/AIDS Rep 2015 Dec;12(4):437-440 [FREE Full text] [doi: 10.1007/s11904-015-0289-1] [Medline: 26412082]
12. Trujillo D, Turner C, Le V, Wilson E, Arayasirikul S. Digital HIV care navigation for young people living with HIV in San Francisco, California: feasibility and acceptability study. JMIR Mhealth Uhealth 2020 Jan 10;8(1):e16838 [FREE Full text] [doi: 10.2196/16838] [Medline: 31922489]
13. Silver L, Johnson C. Internet Connectivity Seen as Having Positive Impact on Life in Sub-Saharan Africa. Pew Research Center. 2018. URL: https://www.pewresearch.org/global/2018/10/09/internet-connectivity-seen-as-having-positive-impact-on-life-in-sub-saharan-africa/ [accessed 2020-11-05]
14. Pellowski JA, Kalichman SC. Recent advances (2011-2012) in technology-delivered interventions for people living with HIV. Curr HIV/AIDS Rep 2012 Dec;9(4):326-334 [FREE Full text] [doi: 10.1007/s11904-012-0133-9] [Medline: 22922945]
15. Daher J, Vijh R, Linthwaite B, Dave S, Kim J, Dheda K, et al. Do digital innovations for HIV and sexually transmitted infections work? Results from a systematic review (1996-2017). BMJ Open 2017 Nov 3;7(11):e017604 [FREE Full text] [doi: 10.1136/bmjopen-2017-017604] [Medline: 29101138]

16. Boulos M, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomed Eng Online 2011 Apr 5;10:24 [FREE Full text] [doi: 10.1186/1475-925X-10-24] [Medline: 21466669]
17. Medhi I, Gautama S, Toyama K. A Comparison of Mobile Money-transfer UIs for Non-literate and Semi-literate Users. In: Conference on Human Factors in Computing Systems. 2009 Presented at: CHI'09; April 4-9, 2009; Boston MA USA. [doi: 10.1145/1518701.1518970]
18. Mathenjwa T, Kim H, Zuma T, Shahmanesh M, Seeley J, Matthews P, et al. Home-based intervention to test and start (HITS) protocol: a cluster-randomized controlled trial to reduce HIV-related mortality in men and HIV incidence in women through increased coverage of HIV treatment. BMC Public Health 2019 Jul 19;19(1):969. [doi: 10.1186/s12889-019-7277-0] [Medline: 31324175]
19. Johnson VD. Promoting behavior change: making healthy choices in wellness and healing choices in illness - use of self-determination theory in nursing practice. Nurs Clin North Am 2007 Jun;42(2):229-41, vi. [doi: 10.1016/j.cnur.2007.02.003] [Medline: 17544680]
20. Deci EL, Ryan RM. The 'What' and 'Why' of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000 Oct;11(4):227-268. [doi: 10.1207/s15327965pli1104_01]
21. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
22. Quinlivan EB, Messer LC, Roytburd K, Blickman A. Unmet core needs for self-determination in HIV-infected women of color in medical care. AIDS Care 2017 May;29(5):603-611 [FREE Full text] [doi: 10.1080/09540121.2016.1243788] [Medline: 27733048]
23. Houston E, McKirnan DJ, Cervone D, Johnson MS, Sandfort TG. Assessing treatment motivation among patients receiving antiretroviral therapy: a multidimensional approach. Psychol Health 2012;27(6):674-687 [FREE Full text] [doi: 10.1080/08870446.2011.618536] [Medline: 21942538]
24. Igreja I, Zuroff DC, Koestner R, Saltaris C, Brouillette M, Lalonde R. Applying self‐determination theory to the prediction of distress and well‐being in gay men with HIV and AIDS1. J Appl Social Pyschol 2000 Apr;30(4):686-706. [doi: 10.1111/j.1559-1816.2000.tb02819.x]
25. Moller AC, Merchant G, Conroy DE, West R, Hekler E, Kugler KC, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. J Behav Med 2017 Feb;40(1):85-98. [doi: 10.1007/s10865-016-9818-7] [Medline: 28058516]
26. 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]
27. Herbst K, Law M, Geldsetzer P, Tanser F, Harling G, Bärnighausen T. Innovations in health and demographic surveillance systems to establish the causal impacts of HIV policies. Curr Opin HIV AIDS 2015 Nov;10(6):483-494 [FREE Full text] [doi: 10.1097/COH.0000000000000203] [Medline: 26371462]
28. Vandormael A, de Oliveira T, Tanser F, Bärnighausen T, Herbeck JT. High percentage of undiagnosed HIV cases within a hyperendemic South African community: a population-based study. J Epidemiol Community Health 2018 Feb;72(2):168-172. [doi: 10.1136/jech-2017-209713] [Medline: 29175867]
29. Barnighausen T, Herbst A, Tanser F, Newell M, Bor J. Unequal Benefits From ART: A Growing Male Disadvantage in Life Expectancy in Rural South Africa. In: The Annual Conference on Retroviruses and Opportunistic Infections. 2014 Presented at: CROI'14; March 3-6, 2014; Boston, Massachusetts.
30. Baisley K, Seeley J, Siedner M, Koole K, Matthews P, Tanser F, et al. Findings from home-based HIV testing and facilitated linkage after scale-up of test and treat in rural South Africa: young people still missing. HIV Med 2019 Nov;20(10):704-708. [doi: 10.1111/hiv.12787] [Medline: 31454139]
31. Cox V, Campbell A, Raphahlelo N, McIntyre J, Rebe K. Towards a male-centered model of care: understanding male preferences and barriers to HIV care in Khayelitsha, South Africa. 2013 Presented at: 7th IAS Conference on HIV Pathogenesis, Treatment and Prevention; 2013; Kuala Lampur.
32. Groh K, Audet CM, Baptista A, Sidat M, Vergara A, Vermund SH, et al. Barriers to antiretroviral therapy adherence in rural Mozambique. BMC Public Health 2011 Aug 16;11:650 [FREE Full text] [doi: 10.1186/1471-2458-11-650] [Medline: 21846344]
33. Heestermans T, Browne JL, Aitken SC, Vervoort SC, Klipstein-Grobusch K. Determinants of adherence to antiretroviral therapy among HIV-positive adults in sub-Saharan Africa: a systematic review. BMJ Glob Health 2016;1(4):e000125 [FREE Full text] [doi: 10.1136/bmjgh-2016-000125] [Medline: 28588979]
34. Fitzgerald M, Collumbien M, Hosegood V. 'No one can ask me 'Why do you take that stuff?': men's experiences of antiretroviral treatment in South Africa. AIDS Care 2010 Mar;22(3):355-360. [doi: 10.1080/09540120903111536] [Medline: 20390516]
35. Cornell M, Cox V, Wilkinson L. Public health blindness towards men in HIV programmes in Africa. Trop Med Int Health 2015 Dec;20(12):1634-1635 [FREE Full text] [doi: 10.1111/tmi.12593] [Medline: 26325025]

36. Bhagwanjee A, Govender K, Akintola O, Petersen I, George G, Johnstone L, et al. Patterns of disclosure and antiretroviral treatment adherence in a South African mining workplace programme and implications for HIV prevention. Afr J AIDS Res 2011;10(Suppl 1):357-368. [doi: 10.2989/16085906.2011.637737] [Medline: 25865512]
37. Mills EJ, Beyrer C, Birungi J, Dybul MR. Engaging men in prevention and care for HIV/AIDS in Africa. PLoS Med 2012 Feb;9(2):e1001167 [FREE Full text] [doi: 10.1371/journal.pmed.1001167] [Medline: 22346735]
38. Kumwenda MK. The influence of masculinity on HIVST community intervention: a qualitative evaluation of empirical evidence from Blantyre, Malawi. Malawi-Liverpool Welcome Trust, Behaviour and Health, Blantyre, Malawi. URL: https://hivstar.lshtm.ac.uk/files/2016/12/The-influence-of-masculinity-on-HIVST-community-intervention.pdf [accessed 2020-11-06]
39. Marson KG, Tapia K, Kohler P, McGrath CJ, John-Stewart GC, Richardson BA, et al. Male, mobile, and moneyed: loss to follow-up vs transfer of care in an urban African antiretroviral treatment clinic. PLoS One 2013;8(10):e78900 [FREE Full text] [doi: 10.1371/journal.pone.0078900] [Medline: 24205345]
40. Chikovore J, Hart G, Kumwenda M, Chipungu GA, Corbett L. 'For a mere cough, men must just chew Conjex, gain strength, and continue working': the provider construction and tuberculosis care-seeking implications in Blantyre, Malawi. Glob Health Action 2015;8:26292 [FREE Full text] [doi: 10.3402/gha.v8.26292] [Medline: 25833138]
41. Hosegood V, Richter L, Clarke L. '...I Should Maintain a Healthy Life Now and Not Just Live as I Please...': men's health and fatherhood in rural South Africa. Am J Mens Health 2016 Nov;10(6):NP39-NP50 [FREE Full text] [doi: 10.1177/1557988315586440] [Medline: 26008734]
42. Siu GE, Wight D, Seeley J. 'Dented' and 'resuscitated' masculinities: the impact of HIV diagnosis and/or enrolment on antiretroviral treatment on masculine identities in rural eastern Uganda. Sahara J 2014;11:211-221 [FREE Full text] [doi: 10.1080/17290376.2014.986516] [Medline: 25444303]
43. Boullé C, Kouanfack C, Laborde-Balen G, Boyer S, Aghokeng AF, Carrieri MP, Stratall ANRS 12110ESTHER Study Group. Gender differences in adherence and response to antiretroviral treatment in the Stratall trial in rural district hospitals in Cameroon. J Acquir Immune Defic Syndr 2015 Jul 01;69(3):355-364. [doi: 10.1097/QAI.0000000000000604] [Medline: 26181708]
44. Ngwenya N, Gumede D, Shahmanesh M, McGrath N, Grant A, Seeley J. Community perceptions of the socio-economic structural context influencing HIV and TB risk, prevention and treatment in a high prevalence area in the era of antiretroviral therapy. Afr J AIDS Res 2018 Mar;17(1):72-81 [FREE Full text] [doi: 10.2989/16085906.2017.1415214] [Medline: 29504507]
45. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
46. User Research Doing Pop-up Research. Government of UK. 2017. URL: https://www.gov.uk/service-manual/user-research/doing-pop-up-research [accessed 2020-11-05]
47. Brooke J. SUS-a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Boca Raton, Florida: CRC Press; 1996.
48. Cordova D, Alers-Rojas F, Lua FM, Bauermeister J, Nurenberg R, Ovadje L, et al. The usability and acceptability of an adolescent mHealth HIV/STI and drug abuse preventive intervention in primary care. Behav Med 2018;44(1):36-47 [FREE Full text] [doi: 10.1080/08964289.2016.1189396] [Medline: 27223646]
49. 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]
50. Entwistle VA, France EF, Wyke S, Jepson R, Hunt K, Ziebland S, et al. How information about other people's personal experiences can help with healthcare decision-making: a qualitative study. Patient Educ Couns 2011 Dec;85(3):e291-e298. [doi: 10.1016/j.pec.2011.05.014] [Medline: 21652162]
51. Lustria ML, Cortese J, Gerend MA, Schmitt K, Kung YM, McLaughlin C. A model of tailoring effects: a randomized controlled trial examining the mechanisms of tailoring in a web-based STD screening intervention. Health Psychol 2016 Nov;35(11):1214-1224. [doi: 10.1037/hea0000399] [Medline: 27441869]
52. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]

Abbreviations

AHRI: Africa Health Research Institute
ART: antiretroviral therapy
ARV: antiretroviral
CAB: community advisory board
CSQ: Client Satisfaction Questionnaire
EPIC-HIV: Empowering people through informed choices for HIV
FGD: focus group discussion

HCI: human-computer interaction
HITS: home-based intervention to test and start
PE: public engagement
SDT: self-determination theory

Edited by G Eysenbach; submitted 22.12.19; peer-reviewed by A Farnham, F Holl, A Rung, C Lancaster; comments to author 07.07.20; revised version received 17.09.20; accepted 22.09.20; published 24.11.20.

Please cite as:
Mathenjwa T, Adeagbo O, Zuma T, Dikgale K, Zeitlin A, Matthews P, Seeley J, Wyke S, Tanser F, Shahmanesh M, Blandford A
Development and Acceptability of a Tablet-Based App to Support Men to Link to HIV Care: Mixed Methods Approach
JMIR Mhealth Uhealth 2020;8(11):e17549
URL: http://mhealth.jmir.org/2020/11/e17549/
doi:10.2196/17549
PMID:33231558

Original Paper

Effect of the Pregnant+ Smartphone App on the Dietary Behavior of Women With Gestational Diabetes Mellitus: Secondary Analysis of a Randomized Controlled Trial

Lisa Garnweidner-Holme[1*], PhD; Lena Henriksen[1*], PhD; Liv Elin Torheim[1*], PhD Prof; Mirjam Lukasse[1,2*], PhD Prof

[1]OsloMet - Oslo Metropolitan University of Applied Sciences, Oslo, Norway

[2]Faculty of Heath and Social Sciences, University of South-Eastern Norway, Campus Vestfold, Norway

[*]all authors contributed equally

Corresponding Author:
Lisa Garnweidner-Holme, PhD
OsloMet - Oslo Metropolitan University of Applied Sciences
St Olavs Plass; PO Box 4
Oslo, 0310
Norway
Phone: 47 48091956
Email: lgarnwei@oslomet.no

Abstract

Background: The prevalence of gestational diabetes mellitus (GDM) is increasing worldwide. A healthy diet and stable blood glucose levels during pregnancy can prevent adverse health outcomes for the mother and the newborn child. Mobile health may be a useful supplement to prenatal care, providing women with targeted dietary information concerning GDM.

Objective: We analyzed secondary data from a two-arm, multicentered, nonblinded randomized controlled trial to determine if a smartphone app with targeted dietary information and blood glucose monitoring had an effect on the dietary behavior of women with GDM.

Methods: Women with a 2-hour oral glucose tolerance test level of ≥9 mmol/L were individually randomized to either the intervention group receiving the Pregnant+ app and usual care or the control group receiving usual care only. Eligible women were enrolled from 5 diabetes outpatient clinics in the Oslo region, Norway, between October 2015 and April 2017. The Pregnant+ app promoted 10 GDM-specific dietary recommendations. A healthy dietary score for Pregnant+ (HDS-P+) was constructed from a 41-item food frequency questionnaire and used to assess the intervention effect on the dietary behavior completed at trial entry and at around gestation week 36. Dietary changes from baseline to week 36 were examined by a paired sample two-tailed *t* test. Between-group dietary differences after the intervention were estimated with analysis of covariance, with adjustment for baseline diet.

Results: A total of 238 women participated: 115 were allocated to the intervention group and 123 to the control group. Of the 238 women, 193 (81.1%) completed the food frequency questionnaire both at baseline and around gestational week 36. All the participants showed improvements in their HDS-P+ from baseline. However, the Pregnant+ app did not have a significant effect on their HDS-P+. The control group reported a higher weekly frequency of choosing fish meals (P=.05). No other significant differences were found between the intervention and control groups. There were no significant demographic baseline differences between the groups, except that more women in the intervention group had a non-Norwegian language as their first language (61 vs 46; P=.02).

Conclusions: Our findings do not support the supplementation of face-to-face follow-up of women with GDM with a smartphone app in the presence of high-standard usual care, as the Pregnant+ app did not have a beneficial effect on pregnant women's diet.

Trial Registration: ClinicalTrials.gov NCT02588729; https://clinicaltrials.gov/ct2/show/NCT02588729

(JMIR Mhealth Uhealth 2020;8(11):e18614) doi:10.2196/18614

KEYWORDS

gestational diabetes mellitus; diet; mHealth; mobile phone; randomized controlled trial

Introduction

Gestational diabetes mellitus (GDM) is defined as hyperglycemia detected at any time during pregnancy [1]. The prevalence of GDM is increasing worldwide and ranges from 1% to 20% globally depending on the screening procedure and population characteristics [2]. The prevalence of GDM in Norway was 5% in 2018, according to the Norwegian Medical Birth Registry [3]. However, a cohort study in a district in Oslo identified GDM in 13% of all women, 11% of ethnic Norwegians, and 12%-17% of women in groups of non-European origin [4]. Women of South Asian and African origins tend to develop GDM at a lower body mass index and age compared to White Europeans [5]. The other risk factors for developing GDM include overweight and obesity, advanced maternal age, a family history of diabetes, and GDM in a previous pregnancy [6]. Even though GDM resolves in most women after delivery, its development may affect the health of both mothers and children in the short and long terms [7,8].

A healthy diet and stable blood glucose levels throughout pregnancy can prevent adverse health outcomes for the mother and the newborn child [9]. About 85% of the women diagnosed with GDM can manage the disease with lifestyle changes such as healthy eating and physical activity, without the need for oral metformin or insulin therapy [10]. However, lifestyle changes presuppose knowledge, motivation, and follow-up by health care professionals [11]. Pregnant women are often in contact with health care professionals; however, perinatal care involves dealing with many health-related issues, and there are some indications that women are not provided sufficient information about the management of GDM by their health care professionals [12,13].

Mobile health (mHealth)—defined as medical and public health practice supported by mobile devices such as smartphones, patient monitoring devices, personal digital assistants, and other wireless devices [14]—may be a useful supplement to perinatal care by providing women with GDM with dietary information and the opportunity to register blood glucose levels [15]. A scoping review has found several ongoing randomized controlled trials (RCTs) that evaluate the effectiveness of smartphone apps in the management of GDM [16]. Some results of these RCTs have been published recently [17,18]. Even though these studies did not find any effect on the glycemic, maternal, and neonatal outcomes [17,18], there is a lack of studies investigating the possible effects of an app on the diet of women with GDM. A systematic review has studied the usability of apps in the health care of pregnant women without GDM. That review indicated that apps may support women in reducing gestational weight gain and in increasing their intake of vegetables and fruits; however, the evidence of their effectiveness is still limited [19]. Dodd et al [20] evaluated the impact of a smartphone app as an adjunct to face-to-face consultation in facilitating dietary changes among pregnant women in South Australia. They found no significant benefit of the smartphone app in the intervention group. All women improved their dietary quality during pregnancy [20].

We have developed the Pregnant+ app for women with GDM [21]. This app provides tailored information on diet, physical activity, breastfeeding, and GDM, and the possibility to automatically transfer or manually record blood glucose levels from a glucometer to the smartphone (Multimedia Appendix 1 and Multimedia Appendix 2). The Pregnant+ app was developed in collaboration with experts in midwifery, obstetrics, physical activity, nutrition, and data security. Pregnant women with GDM of different ethnic origins were involved in several steps of its development [21]. A narrative review on studies with pregnancy-related apps found only 2 multilingual apps for use in prenatal care [22]. The Pregnant+ app is available in Norwegian, Urdu, and Somali languages. Information and pictures related to diet and physical activity are culturally adjusted according to the chosen language. This app was found to be the only app adapted to specific target groups (women born in Pakistan and Somalia) in a scoping review by Chen and Carbone [16]. The effect of this app on the main outcome (2-hour glucose level of the routine postpartum oral glucose tolerance test [OGTT]) was tested in a two-arm RCT at 5 diabetes outpatient clinics in the Oslo region, Norway [18,23]. The study showed that the Pregnant+ app did not have any significant effect on the main outcome [18].

The aim of this study was to examine if the Pregnant+ app had an effect on the dietary behavior of women. No specific dietary recommendation for women with GDM existed when the study was started in 2014. Women with GDM were recommended to follow the national dietary guidelines for healthy eating [24]. Some hospitals developed adjusted dietary advice for pregnant women with GDM. The Pregnant+ app promoted 10 GDM-specific dietary recommendations that were developed in cooperation with clinical nutritionists [21].

Methods

Study Design

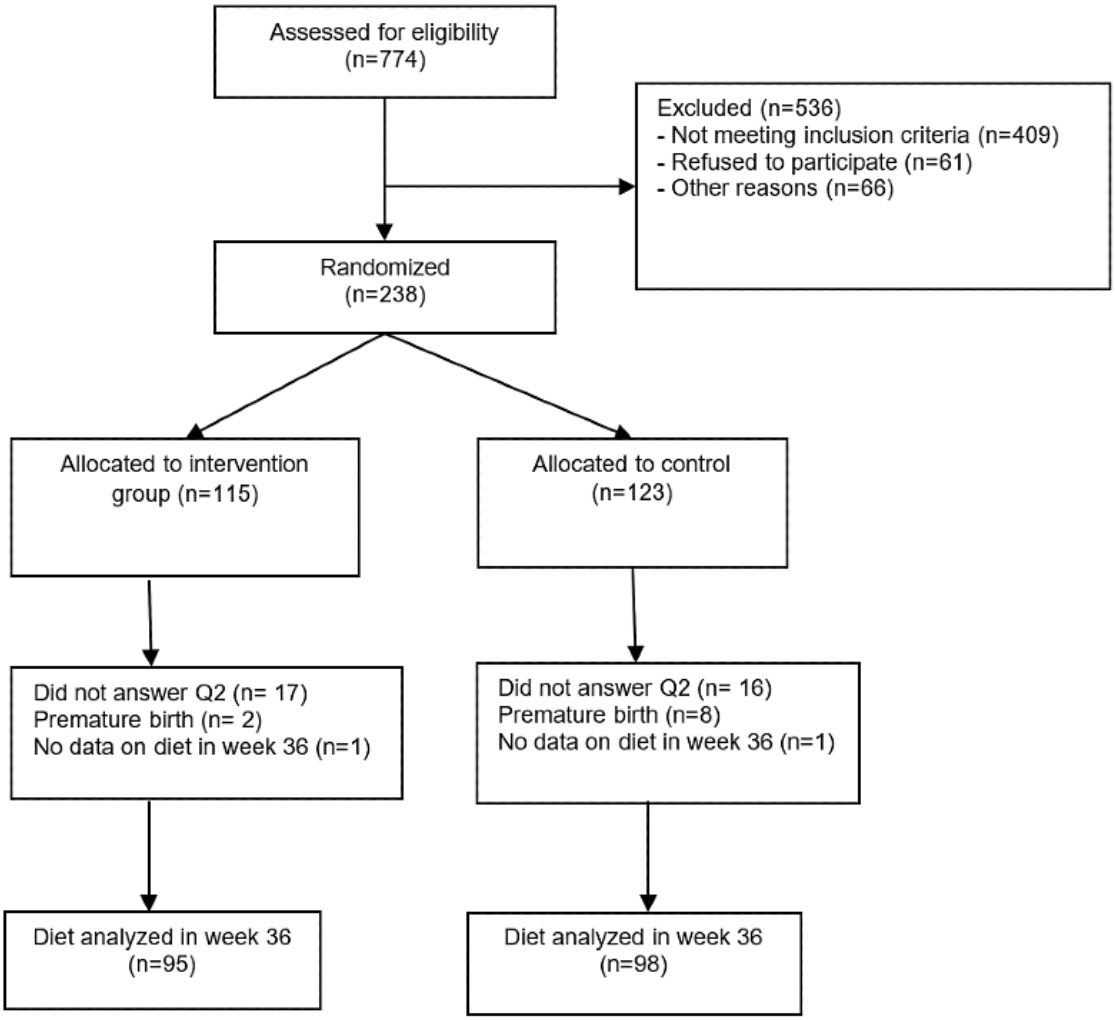
We analyzed secondary data from a two-arm, multicentered, nonblinded RCT for women with GDM, which was conducted at 5 diabetes outpatient clinics in the Oslo region. This RCT is in accordance with the CONSORT-EHEALTH checklist (Multimedia Appendix 3).

Recruitment

Women with GDM were recruited from October 2015 to April 2017 by health care professionals at the diabetes outpatient clinics. At the time of recruitment, pregnant women with higher risk for GDM based on their prepregnancy weight, family history of diabetes, age, and ethnicity were sent for an OGTT [25]. Eligible women for this study were diagnosed with GDM by a 2-hour OGTT blood glucose level of ≥9 mmol/L, according to the definition of GDM in the Norwegian guidelines [25]. In addition, participants were older than 18 years, were less than 33-weeks pregnant, owned smartphones, and understood Norwegian, Urdu, or Somali. Women with type 1 or type 2 diabetes (OGTT blood glucose levels≥11 mmol/L), twin pregnancy, celiac disease, or lactose intolerance were excluded from the study. In total, 774 women were assessed for eligibility and 238 participated (Figure 1). Those who agreed to participate signed a consent form. All the participants in both groups

received a glucometer and lancets from the study administrators. The Norwegian Social Science Data Services (identifier: 2014/38942) approved the study.

Figure 1. Flow chart describing the process leading up to the final number included in the analysis of dietary behavior in the Pregnant+ study.



Randomization and Blinding

The participants were randomly allocated to 2 groups: intervention (access to the Pregnant+ app and usual care) and control (usual care). Randomization was performed on a 1:1 basis with allocated blocks of 4. Women who agreed to participate filled out a baseline questionnaire (Q1) on an electronic tablet (average time 30-45 minutes). After completing Q1, a computer-based program randomized and allocated the women to either the intervention or the control group. The participating women, project workers, and health care professionals at the diabetes outpatient clinics were not blinded to the allocation.

Intervention and Control

Usual Care (Control) Group

The participants in the control group received usual care for GDM according to the national guidelines [26]. This included regular consultations (every 1-2 weeks) with midwives or nurses, both specialized in diabetes, at the diabetes outpatient clinics. According to the guidelines [26], women should be provided with information about a healthy diet, with emphasis on regular meals with limited intake of sugar-rich foods and increased intake of whole grains and vegetables. The women in the control group were instructed how to measure their blood glucose levels and were asked to record these levels in a paper diary. They received written and verbal dietary advice on the basis of their blood glucose levels. In the absence of specific dietary guidelines for women with GDM, the different diabetes outpatient clinics included in the study developed some specific dietary guidelines for these women that emphasized regular meals, increased intake of vegetables and whole grains, and limited intake of sugar. If women in the usual care group downloaded the Pregnant+ app, their access was restricted to a single page with a link to the website of the Norwegian Directorate of Health with generic health information for women with GDM and a link to the Norwegian Federation of Diabetes.

Pregnant+ App and Usual Care (Intervention) Group

The participants in the intervention group had access to the Pregnant+ app in addition to usual care, as described above. The women allocated to the app group could download the app from the Apple Store or Google Play at the hospital or home. The app contained 4 main icons: “Blood glucose,” “Physical

activity," "Food and beverages," and "Diabetes information." The 10 GDM-specific dietary recommendations that were developed for this study [21] were presented in the "Food and beverages" icon for women (Textbox 1). The women could select if the dietary recommendations should be presented with food items and pictures representing the Norwegian, Urdu, or Somali food culture. They were also referred to recipes on the home page from the Norwegian Diabetes Foundation. They could automatically transfer or manually register their blood glucose levels in the icon "Blood glucose." After registering, they received feedback on their values. Those with too high values were directly referred to the dietary recommendations. A graphical representation of the blood glucose levels visually aided the monitoring.

Textbox 1. The dietary recommendations in the Pregnant+ app.

- Eat healthy meals regularly.
- Eat and drink little sugar.
- Eat more vegetables.
- Choose whole-grain products.
- Limit your intake of salt.
- Eat enough fish.
- Choose lean dairy milk produce.
- Choose healthy and less oil.
- Read nutrition labels on foods before buying.
- Choose water when thirsty.

Measurements

The participants answered the questionnaires on an electronic tablet during their first consultation at a diabetes outpatient clinic and at their consultation around gestational week 36. The questionnaire included a 41-item food frequency questionnaire (FFQ). At baseline, they were asked to report their dietary habits prior to being diagnosed with GDM. In the second questionnaire, they were asked to report their current diet. The FFQ included the following food groups: beverages, milk and dairy products, bread and grain, fruit and vegetables, snacks, meat, and ready-to-eat meals. Answers to the questions on the frequency of intake ranged from 0 (never) to 9 (several times daily). The FFQ was based on the Fit for Delivery study and has been shown to have an adequate level of test-retest reliability [26]. The FFQs in Somali and Urdu were tested for comprehension and appropriateness by conducting qualitative interviews with Somali and Pakistani Norwegian women.

The healthy diet score for Pregnant+ (HDS-P+) was constructed using 9 subscales, with a possible range of 0 to 90. The subscales were constructed on the basis of the dietary recommendations (second point to tenth point) provided in Textbox 1 and consisted of different questions in the FFQ related to the dietary recommendations. The women were asked how often they choose different food groups, with the following answer options: 0=never, 1=less than once a week, 2=once a week, 3=twice, 4=three times, 5=four times, 6=five times, 7=six times, 8=every day, and 9=several times a day.

Information on background characteristics was obtained from the baseline questionnaire and consisted of different socioeconomic variables: age, education, income, country of birth, marital status, economic hardship, and language. Other variables related to pregnancy and health were parity, gestational age at baseline, prior GDM, and perceived health score [23].

Statistical Analysis

Maternal baseline characteristics were compared according to randomization status. The characteristics were presented as mean (SD) for continuous variables (independent sample two-tailed *t* test) and proportions (%) for categorical variables (χ^2 test). Dietary changes from baseline to around gestational week 36 were examined by a paired sample two-tailed *t* test. A one-way between-group analysis (analysis of covariance) was conducted to measure the effect of the Pregnancy+ app. The dietary behavior after the intervention was examined with adjustment to the baseline values. The HDS-P+ and subscales related to the recommended dietary advice were the dependent variables, and the randomization status (use of the Pregnant+ app or not) was the independent variable. Sensitivity analysis was performed to evaluate the effect of the differences in nonnative and native Norwegian-speaking women between the intervention and control groups at baseline. This did not alter the results and the final model did not adjust for this. Levene test and normality checks were carried out and the assumptions were met. All statistical analyses were performed with SPSS for IBM statistical software package (version 25, IBM Corporation). A two-sided *P* value of ≤.05 was considered significant.

Power

The power calculation was for the primary outcome for the RCT [23].

Data Exclusion

Figure 1 presents the flowchart for this study. Two women were excluded because of missing dietary data at gestational week 36. No other participant had more than 2 values missing in the 41-item FFQ. The missing data in this study were not imputed.

Results

Participant Characteristics

A total of 238 women were recruited at 5 diabetes outpatient clinics in the southeast region of Norway and randomized to use the Pregnant+ app (intervention group, n=115) or no app (control group, n=123). Figure 1 shows an overview of the final numbers in the dietary analysis. Of the 238 women, 193 (81.1%) women completed the FFQ both at baseline and at gestational week 36. Background characteristics are described according to the randomization status (Table 1). There were no significant baseline differences between the groups, except for more women with a non-Norwegian language as their first language in the intervention group (61 vs 46, *P*=.02).

Table 1. Background characteristics at baseline of the participants who provided dietary data at baseline (Q1) and after the intervention (Q2) in the study on the Pregnant+ app.

Background characteristics	Total, N=193	Control group, n=98	Intervention group, n=95	*P* value
Age (years), n (%)				.11
≤29	47 (24.4)	22 (22)	25 (26)	
30-37	110 (57.0)	52 (53)	58 (61)	
≥38	36 (18.7)	24 (25)	12 (13)	
Gestational age at baseline, mean (SD)	27.1 (4.6)	27.3 (4.6)	26.9 (4.5)	.66
Parity, n (%)				.21
Primiparous	86 (44.6)	48 (49)	38 (40)	
Multiparous	107 (55.4)	50 (51)	57 (60)	
Previous GDM[a] (N=107[b]), n (%)				.82
No	75(70.1)	34 (68)	41 (72)	
Yes	32 (29.9)	16 (32)	16 (28)	
BMI (N=190[c]), n (%)				.68
<24.9	83 (43.7)	44 (45)	39 (42)	
25.0-29.9	57 (30.0)	26 (27)	31 (33)	
30.0-34.9	31 (16.3)	18 (19)	13 (14)	
35.0-45.0	19 (10.0)	9 (9)	10 (11)	
Country of birth, n (%)				.15
Norway	90 (46.6)	52 (53)	38 (40)	
Western Europe + United States of America	13 (6.7)	9 (9)	4 (4)	
Eastern Europe	18 (9.3)	9 (9)	9 (10)	
Asia	45 (23.3)	16 (16)	29 (31)	
Africa	22 (11.4)	10 (10)	12 (13)	
South America	5 (2.6)	2 (2)	3 (3)	
Marital status, n (%)				.62
Married/cohabiting	179 (92.7)	90 (92)	89 (94)	
Single/other	14 (7.3)	8 (8)	6 (6)	
Education, n (%)				.51
Primary school/no education	19 (9.8)	12 (12)	7 (7)	
High school	40 (20.7)	23 (24)	17 (18)	
College/university<4 years	47 (24.4)	23 (24)	24 (25)	
College/university≥4 years	87 (45.1)	40 (41)	47 (50)	
Smoking or wet tobacco , n (%)				.69
No	189 (97.9)	96 (98)	93 (98)	
Yes	4 (2.1)	2 (2)	2 (2)	
Main activity, n (%)				.22
Employed or self-employed	147 (76.2)	71 (72)	76 (80)	
Not employed or not self-employed	46 (23.8)	27 (28)	19 (20)	
Joined income, n (%)				.78
≤59,900 USD	57 (29.9)	26 (27)	31 (33)	
60,000-79,900 USD	28 (14.2)	14 (14)	14 (15)	
80,000-99,900 USD	39 (19.8)	20 (20)	19 (20)	

Background characteristics	Total, N=193	Control group, n=98	Intervention group, n=95	*P* value
≥100,000 USD	35 (18.8)	20 (20)	15 (16)	
I don't know	34 (17.3)	18 (18)	16 (17)	
Economic hardship (N=188[c]), n (%)				.85
No	58 (30.9)	29 (30)	29 (32)	
Yes	130 (69.1)	67 (70)	63 (69)	
Language, n (%)				.02
Native Norwegian-speaking	86 (45.1)	52 (53)	34 (36)	
Nonnative Norwegian-speaking	107 (54.9)	46 (47)	61 (64)	
Perceived health score (0-100), mean (SD)	70.8 (19.7)	70.5 (20.5)	71.2 (18.9)	.80

[a]GDM: gestational diabetes mellitus.

[b]Among multiparous women only.

[c]Some values are missing.

Outcomes

Dietary Outcomes Around Gestational Week 36

Overall, the total HDS-P+ and most of the subscales, except the intake of healthy oils, improved from baseline to gestational week 36 (Table 2).

Table 2. Dietary changes from baseline to gestational week 36.[a]

Subscales	Baseline values, mean (SD)	Week 36 values, mean (SD)	*P* value
1. HDS-P+[b]	40.36 (14.11)	55.56 (13.70)	<.001
2. Sugar (times/week)	10.10 (7.88)	1.89 (3.21)	<.001
3. Vegetables (times/week)	8.87 (3.52)	10.35 (3.5)	<.001
4. Whole grains (times/week)	6.71 (2.96)	8.87 (2.78)	<.001
5. Salt (times/week)	3.71 (3.10)	2.39 (2.49)	<.001
6. Fish (times/week)	1.84 (1.17)	2.21 (1.32)	<.001
7. Low-fat milk (times/week)	4.84 (4.08)	4.22 (3.34)	.02
8. Healthy oil (% of total dietary fat)[c]	62.41 (25.32)	65.09 (25.36)	.11
9. Read nutrition labels	5.78 (3.44)	8.45 (2.64)	<.001
10. Water (% of total fluid intake)[c]	40.15 (14.67)	51.21 (17.35)	<.001

[a]Paired sample two-tailed *t* test.

[b]HDS-P+: healthy dietary score for Pregnant+.

[c]Percentage of weekly consumption.

Between-Group Differences After Intervention

A one-way between-group analysis of covariance was conducted to compare the effectiveness of the app on the participants' dietary habits after being diagnosed with GDM. The women's HDS-P+ preintervention was used as the covariate in the analysis. Table 3 presents the between-group differences for the overall HDS-P+ and the 9 different subscales at gestational week 36. No significant differences favored the intervention group. The analysis showed that the control group reported to eat more fish meals per week (*P*=.05). No other significant differences were found between the intervention and control groups.

Table 3. Between-group differences in 10 dietary domains reported after the intervention (gestational week 36) in the Pregnant+ app.

Dietary domain	Subscale(s)	Control group, n=98, mean (SE)	Intervention group, n=95, mean (SE)	Estimated difference after intervention[a], mean (SE)	95% CI	*P* value
Eat healthy	HDS-P+[b]	56.11 (1.11)	55.34 (1.13)	0.77 (1.72)	–2.62, 4.16	.65
Eat and drink little sugar	Sugar (times/week)	1.97 (0.31)	1.79 (0.32)	0.18 (0.45)	–0.70, 1.06	.68
Eat more vegetables	Vegetables (times/week)	10.40 (0.32)	10.30 (0.32)	0.09 (0.45)	–0.83, 0.95	.86
Choose whole grains	Whole grains (times/week)	8.23 (0.27)	8.01 (0.28)	0.14 (0.40)	–0.79, 0.99	.73
Limit your intake of salt	Salt (times/week)	2.53 (0.12)	2.25 (0.21)	0.28 (0.31)	–0.31, 0.88	.35
Eat enough fish	Fish (times/week)	2.34 (0.09)	2.09 (0.09)	0.26 (0.13)	–0.01, 0.51	.05
Choose lean dairy milk	Low-fat milk (times/week)	4.40 (0.28)	4.04 (2.85)	0.35 (0.40)	–0.42, 1.14	.38
Eat less saturated fat	Healthy oil (% of total dietary fat)[c]	66.30 (2.05)	63.80 (2.10)	2.50 (2.50)	–3.31, 8.33	.40
Read nutrition labels	Read labels	8.76 (0.25)	8.13 (0.25)	0.63 (0.36)	–0.07, 1.33	.08
Choose water	Water (% of total fluid intake) [c]	51.93 (1.57)	50.47 (1.59)	1.46 (2.24)	–2.69, 5.89	.57

[a]Analysis of covariance adjusted for baseline HDS-P+.

[b]HDS-P+: healthy dietary score for Pregnant+.

[c]Percentage of weekly consumption.

Discussion

Principal Results

The Pregnant+ app combined with usual care did not have any significant effect on the dietary behavior of the participants during pregnancy compared to the dietary behavior of the participants receiving usual care only. All the participants improved their diet from the time they were diagnosed with GDM to gestational week 36.

Comparison With Prior Work

This study adds to the literature on the development and effect of pregnancy-related apps for the management of GDM and for following a healthy diet [22,27,28]. Pregnant women consider these apps to be useful and convenient for nutrition information and management of their diets [27]; however, little is known about their effects on the dietary behavior [19]. mHealth apps may provide several functions targeting behavior change or monitoring. The most successful smartphone-based interventions for dietary change and health outcomes include elements of self-monitoring and personalized feedback [29]. Similar to other apps for women with GDM [21,30], the Pregnant+ app includes a function for self-monitoring of blood glucose levels. According to our qualitative study on women's experience with the Pregnant+ app, the self-management of blood glucose levels was the most important aspect of the app for increasing self-awareness and motivation [31]. Ten of the 17 participants from the intervention group reported to use the Pregnant+ app daily for their blood glucose management. However, the monitoring of food intake was not possible in the app. A qualitative study about the acceptability of a smartphone app for patients with type 2 diabetes indicated that the use of a digital diabetes diary to monitor food intake supported them in eating a healthy diet [32]. Dodd et al [20] assessed the effect of a smartphone app on the dietary behavior of pregnant women. Their app included a combination of information provision, goal setting, feedback, and self-monitoring. The use of the app was poor, and it provided no additional benefit over face-to-face consultation and printed materials in improving dietary behaviors [20]. It should also be considered that adherence to self-monitoring has been shown to decrease over time and that self-monitoring is successful only when people are regularly reminded to use the app [33]. The women in our intervention group were not reminded to use the app.

Compared to studies demonstrating the positive impact of apps on healthy eating and blood glucose levels [29,34,35], the Pregnant+ app did not have any personal interaction with the women. For instance, a review on the use of telemedicine technology for managing diabetes in pregnancy (not just GDM) showed a modest but statistically significant improvement in HbA_{1c} levels [35]. To avoid bias, the health care professionals were not asked to use the app actively during their consultations. However, the women automatically received specific dietary information when registering too high blood glucose levels in the Pregnant+ app. The women in our qualitative study wanted more involvement of the health care professionals in the usage of the app. In an RCT involving 203 pregnant women in the United Kingdom, women with access to a smartphone app had a higher level of satisfaction with care than the women in the control group [17]. In this study, the midwives checked the women's registered blood glucose levels in the app 3 times a week and sent feedback via SMS text messaging [17].

Similar to that reported in other mHealth studies [36], technical problems with the app could be a reason for us to not find any effect of the app on the participants' diets. Some of them experienced problems with the automatic transfer of the blood glucose levels [31]. A cross-sectional survey on the use of mHealth among Latino patients with diabetes found that the lack of operability between the smartphone app and other devices could serve as a barrier to using the app [37].

All the participants in this study improved their diet after being diagnosed with GDM. This is in accordance with previous research, indicating that the diagnosis of GDM motivates women to change their diets [38-40]. Our previous study about women's dietary habits prior to being diagnosed with GDM showed low adherence to national dietary recommendations [41]. A significantly higher proportion of nonnative Norwegian-speaking women had a high healthy diet score compared with native Norwegian-speaking women. In this study, significantly more women with a non-Norwegian language as their first language were in the intervention group (P=.02). Previous research has shown that a combination of technological, health literacy, and language issues may result in a lower uptake of pregnancy apps among immigrant women [16]. These barriers may also have led to a lower usage of the Pregnant+ app.

Limitations

One of the main limitations of this study was that we did not have access to usage logs because of technical problems. To secure the participants' privacy, we did not collect any additional data from the app. We do not know if those in the intervention group actually used the Pregnant+ app or about their frequency of usage or the pages in the app accessed by them. Our qualitative study on 17 participants from the intervention group showed that some women used the app regularly and some did not use it at all because of technical problems [31]. Patients participating in a study will often experience an effect even when not receiving the intervention—the Hawthorne effect. In our study, this could account for the lack of effects as women not using the app may have focused more on a healthy diet as a result of participating in a study. The data for this study were derived from self-completed questionnaires, which include the possibility for recall bias. Social desirability might have biased the self-report of dietary intake. The FFQ covered only selected aspects of the overall diet. Thus, the difference in the HDS-P+ should not be interpreted as an absolute measure of dietary change [42].

Conclusions

To our knowledge, this is one of the first studies to evaluate the effect of a smartphone app on the dietary behavior of women with GDM [17]. Our findings do not support the supplementation of face-to-face follow-up of women with GDM with a smartphone app in the presence of high-standard usual care. However, the app might be a useful tool for women who do not receive sufficient dietary counselling in person. Future research should explore the effects of various technological features provided in a smartphone app to improve the care of women with GDM.

Acknowledgments

This study was funded by the Norwegian Research Council (identifier: 228517). We also acknowledge the participating women as well as the health care professionals and project members involved in the recruitment process.

Authors' Contributions

LGH wrote the manuscript. LH and LET conducted all statistical analysis. All authors planned the study and reviewed the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshot of the app.
[PNG File , 309 KB - mhealth_v8i11e18614_app1.png]

Multimedia Appendix 2
Screenshot of the dietary recommendation in the app.
[PNG File , 212 KB - mhealth_v8i11e18614_app2.png]

Multimedia Appendix 3
CONSORT-EHEALTH Checklist (V 1.6.1).
[PDF File (Adobe PDF File), 359 KB - mhealth_v8i11e18614_app3.pdf]

References

1. The WHO. WHO recommendation on the diagnosis of gestational diabetes in pregnancy. Geneva, Switzerland: The WHO Reproductive Health Library; 2016. URL: https://tinyurl.com/y4lprrxw [accessed 2019-12-04]

2. Rani PR, Begum J. Screening and Diagnosis of Gestational Diabetes Mellitus, Where Do We Stand. J Clin Diagn Res 2016 Apr;10(4):QE01-QE04 [FREE Full text] [doi: 10.7860/JCDR/2016/17588.7689] [Medline: 27190902]
3. Medical Birth Registry Norway, The Medical Birth Registry Norway. 2018. URL: https://www.fhi.no/hn/helseregistre-og-registre/mfr/ [accessed 2019-12-04]
4. Jenum AK, Diep LM, Holmboe-Ottesen G, Holme IMK, Kumar BN, Birkeland KI. Diabetes susceptibility in ethnic minority groups from Turkey, Vietnam, Sri Lanka and Pakistan compared with Norwegians - the association with adiposity is strongest for ethnic minority women. BMC Public Health 2012 Mar 01;12:150 [FREE Full text] [doi: 10.1186/1471-2458-12-150] [Medline: 22380873]
5. Makgoba M, Savvidou MD, Steer PJ. An analysis of the interrelationship between maternal age, body mass index and racial origin in the development of gestational diabetes mellitus. BJOG 2012 Feb;119(3):276-282 [FREE Full text] [doi: 10.1111/j.1471-0528.2011.03156.x] [Medline: 22044452]
6. Schneider S, Hoeft B, Freerksen N, Fischer B, Roehrig S, Yamamoto S, et al. Neonatal complications and risk factors among women with gestational diabetes mellitus. Acta Obstet Gynecol Scand 2011 Mar;90(3):231-237 [FREE Full text] [doi: 10.1111/j.1600-0412.2010.01040.x] [Medline: 21306307]
7. Aziz S, Munim TF, Fatima SS. Post-partum follow-up of women with gestational diabetes mellitus: effectiveness, determinants, and barriers. J Matern Fetal Neonatal Med 2018 Jun;31(12):1607-1612. [doi: 10.1080/14767058.2017.1321630] [Medline: 28423981]
8. Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The Pathophysiology of Gestational Diabetes Mellitus. Int J Mol Sci 2018 Oct 26;19(11) [FREE Full text] [doi: 10.3390/ijms19113342] [Medline: 30373146]
9. Landon MB, Spong CY, Thom E, Carpenter MW, Ramin SM, Casey B, Eunice Kennedy Shriver National Institute of Child HealthHuman Development Maternal-Fetal Medicine Units Network. A multicenter, randomized trial of treatment for mild gestational diabetes. N Engl J Med 2009 Oct 01;361(14):1339-1348 [FREE Full text] [doi: 10.1056/NEJMoa0902430] [Medline: 19797280]
10. American Diabetes Association. Management of Diabetes in Pregnancy. Diabetes Care 2017;40(Suppl 1):S114-S119. [doi: 10.2337/dc17-S016] [Medline: 27979900]
11. Schiavo R. Health communication: from theory to practice. 1st ed. San Francisco, CA: Jossey-Bass; 2007.
12. Carolan M, Gill GK, Steele C. Women's experiences of factors that facilitate or inhibit gestational diabetes self-management. BMC Pregnancy Childbirth 2012 Sep 18;12:99 [FREE Full text] [doi: 10.1186/1471-2393-12-99] [Medline: 22988897]
13. Borgen I, Garnweidner-Holme LM, Jacobsen AF, Fayyad S, Cvancarova Småstuen M, Lukasse M. Knowledge of gestational diabetes mellitus at first consultation in a multi-ethnic pregnant population in the Oslo region, Norway - a cross-sectional study. Ethn Health 2019 Aug 16:1-14. [doi: 10.1080/13557858.2019.1655530] [Medline: 31416352]
14. Frequently asked questions on global task force on digital health for TB and its work. World Health Organization. URL: http://www.who.int/tb/areas-of-work/digital-health/faq/en/ [accessed 2020-09-09]
15. Chan KL, Chen M. Effects of Social Media and Mobile Health Apps on Pregnancy Care: Meta-Analysis. JMIR Mhealth Uhealth 2019 Jan 30;7(1):e11836 [FREE Full text] [doi: 10.2196/11836] [Medline: 30698533]
16. Chen Q, Carbone ET. Functionality, Implementation, Impact, and the Role of Health Literacy in Mobile Phone Apps for Gestational Diabetes: Scoping Review. JMIR Diabetes 2017 Oct 04;2(2):e25 [FREE Full text] [doi: 10.2196/diabetes.8045] [Medline: 30291088]
17. Mackillop L, Hirst JE, Bartlett KJ, Birks JS, Clifton L, Farmer AJ, et al. Comparing the Efficacy of a Mobile Phone-Based Blood Glucose Management System With Standard Clinic Care in Women With Gestational Diabetes: Randomized Controlled Trial. JMIR Mhealth Uhealth 2018 Mar 20;6(3):e71 [FREE Full text] [doi: 10.2196/mhealth.9512] [Medline: 29559428]
18. Borgen I, Småstuen MC, Jacobsen AF, Garnweidner-Holme LM, Fayyad S, Noll J, et al. Effect of the Pregnant+ smartphone application in women with gestational diabetes mellitus: a randomised controlled trial in Norway. BMJ Open 2019 Nov 11;9(11):e030884. [doi: 10.1136/bmjopen-2019-030884] [Medline: 31719080]
19. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The Usability and Effectiveness of Mobile Health Technology-Based Lifestyle and Medical Intervention Apps Supporting Health Care During Pregnancy: Systematic Review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
20. Dodd JM, Louise J, Cramp C, Grivell RM, Moran LJ, Deussen AR. Evaluation of a smartphone nutrition and physical activity application to provide lifestyle advice to pregnant women: The SNAPP randomised trial. Matern Child Nutr 2018 Jan;14(1) [FREE Full text] [doi: 10.1111/mcn.12502] [Medline: 28836373]
21. Garnweidner-Holme LM, Borgen I, Garitano I, Noll J, Lukasse M. Designing and Developing a Mobile Smartphone Application for Women with Gestational Diabetes Mellitus Followed-Up at Diabetes Outpatient Clinics in Norway. Healthcare (Basel) 2015 May 21;3(2):310-323 [FREE Full text] [doi: 10.3390/healthcare3020310] [Medline: 27417764]
22. Hughson JP, Daly JO, Woodward-Kron R, Hajek J, Story D. The Rise of Pregnancy Apps and the Implications for Culturally and Linguistically Diverse Women: Narrative Review. JMIR Mhealth Uhealth 2018 Nov 16;6(11):e189 [FREE Full text] [doi: 10.2196/mhealth.9119] [Medline: 30446483]

23. Borgen I, Garnweidner-Holme LM, Jacobsen AF, Bjerkan K, Fayyad S, Joranger P, et al. Smartphone application for women with gestational diabetes mellitus: a study protocol for a multicentre randomised controlled trial. BMJ Open 2017 Mar 27;7(3):e013117. [doi: 10.1136/bmjopen-2016-013117] [Medline: 28348183]
24. Holan S, Matheison M, Petersen K. A national clinical guideline for antenatal care. Short version. Norway, Oslo: Directorate for Health and Social Affairs; 2005.
25. Henriksen T, Thordardson H, Strøm-Raoum E. Svangerskapsdiabetes. Website. 2014. URL: http://legeforeningen.no/Fagmed/Norsk-gynekologisk-forening/Veiledere/Veileder-ifodselshjelp-2014/Diabetes-i-svangerskapet/ [accessed 2019-12-13]
26. Øverby NC, Hillesund ER, Sagedal LR, Vistad I, Bere E. The Fit for Delivery study: rationale for the recommendations and test-retest reliability of a dietary score measuring adherence to 10 specific recommendations for prevention of excessive weight gain during pregnancy. Matern Child Nutr 2015 Jan;11(1):20-32 [FREE Full text] [doi: 10.1111/mcn.12026] [Medline: 23241065]
27. Wang N, Deng Z, Wen LM, Ding Y, He G. Understanding the Use of Smartphone Apps for Health Information Among Pregnant Chinese Women: Mixed Methods Study. JMIR Mhealth Uhealth 2019 Jun 18;7(6):e12631 [FREE Full text] [doi: 10.2196/12631] [Medline: 31215516]
28. Lupton D, Pedersen S. An Australian survey of women's use of pregnancy and parenting apps. Women Birth 2016 Aug;29(4):368-375. [doi: 10.1016/j.wombi.2016.01.008] [Medline: 26874938]
29. Kankanhalli A, Shin J, Oh H. Mobile-Based Interventions for Dietary Behavior Change and Health Outcomes: Scoping Review. JMIR Mhealth Uhealth 2019 Jan 21;7(1):e11312 [FREE Full text] [doi: 10.2196/11312] [Medline: 30664461]
30. Hirst JE, Mackillop L, Loerup L, Kevat DA, Bartlett K, Gibson O, et al. Acceptability and user satisfaction of a smartphone-based, interactive blood glucose management system in women with gestational diabetes mellitus. J Diabetes Sci Technol 2015 Jan;9(1):111-115. [doi: 10.1177/1932296814556506] [Medline: 25361643]
31. Skar JB, Garnweidner-Holme LM, Lukasse M, Terragni L. Women's experiences with using a smartphone app (the Pregnant+ app) to manage gestational diabetes mellitus in a randomised controlled trial. Midwifery 2018 Mar;58:102-108. [doi: 10.1016/j.midw.2017.12.021] [Medline: 29329023]
32. Torbjørnsen A, Ribu L, Rønnevig M, Grøttland A, Helseth S. Users' acceptability of a mobile application for persons with type 2 diabetes: a qualitative study. BMC Health Serv Res 2019 Sep 06;19(1):641 [FREE Full text] [doi: 10.1186/s12913-019-4486-2] [Medline: 31492176]
33. McCarroll R, Eyles H, Ni Mhurchu C. Effectiveness of mobile health (mHealth) interventions for promoting healthy eating in adults: A systematic review. Prev Med 2017 Dec;105:156-168. [doi: 10.1016/j.ypmed.2017.08.022] [Medline: 28882743]
34. Radhakrishnan K. The efficacy of tailored interventions for self-management outcomes of type 2 diabetes, hypertension or heart disease: a systematic review. J Adv Nurs 2012 Mar;68(3):496-510. [doi: 10.1111/j.1365-2648.2011.05860.x] [Medline: 22010971]
35. 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]
36. Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Patient Educ Couns 2010 Mar;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]
37. Arora S, Ford K, Terp S, Abramson T, Ruiz R, Camilon M, et al. Describing the evolution of mobile technology usage for Latino patients and comparing findings to national mHealth estimates. J Am Med Inform Assoc 2016 Sep;23(5):979-983 [FREE Full text] [doi: 10.1093/jamia/ocv203] [Medline: 26995564]
38. Atakora L, Poston L, Hayes L, Flynn AC, White SL. Influence of GDM Diagnosis and Treatment on Weight Gain, Dietary Intake and Physical Activity in Pregnant Women with Obesity: Secondary Analysis of the UPBEAT Study. Nutrients 2020 Jan 30;12(2) [FREE Full text] [doi: 10.3390/nu12020359] [Medline: 32019123]
39. Jelsma JGM, van Leeuwen KM, Oostdam N, Bunn C, Simmons D, Desoye G, et al. Beliefs, Barriers, and Preferences of European Overweight Women to Adopt a Healthier Lifestyle in Pregnancy to Minimize Risk of Developing Gestational Diabetes Mellitus: An Explorative Study. J Pregnancy 2016;2016:3435791 [FREE Full text] [doi: 10.1155/2016/3435791] [Medline: 26885396]
40. Jones EJ, Roche CC, Appel SJ. A review of the health beliefs and lifestyle behaviors of women with previous gestational diabetes. J Obstet Gynecol Neonatal Nurs 2009;38(5):516-526. [doi: 10.1111/j.1552-6909.2009.01051.x] [Medline: 19883473]
41. Garnweidner-Holme L, Torheim LE, Henriksen L, Borgen I, Holmelid S, Lukasse M. Adherence to the Norwegian dietary recommendations in a multi-ethnic pregnant population prior to being diagnosed with gestational diabetes mellitus. Food Sci Nutr 2020 Jul;8(7):3031-3040 [FREE Full text] [doi: 10.1002/fsn3.1248] [Medline: 32724567]
42. Hillesund ER, Bere E, Sagedal LR, Vistad I, Øverby NC. Effect of a diet intervention during pregnancy on dietary behavior in the randomized controlled Norwegian Fit for Delivery study. J Dev Orig Health Dis 2016 Oct;7(5):538-547. [doi: 10.1017/S2040174416000258] [Medline: 27307037]

Abbreviations

FFQ: food frequency questionnaire
GDM: gestational diabetes mellitus
HDS-P+: healthy diet score for Pregnant+
mHealth: mobile health
OGTT: oral glucose tolerance test
RCT: randomized controlled trial

Edited by L Buis; submitted 08.03.20; peer-reviewed by LJ Li, A Rosman, E Burner; comments to author 20.04.20; revised version received 23.04.20; accepted 25.09.20; published 04.11.20.

Please cite as:
Garnweidner-Holme L, Henriksen L, Torheim LE, Lukasse M
Effect of the Pregnant+ Smartphone App on the Dietary Behavior of Women With Gestational Diabetes Mellitus: Secondary Analysis of a Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(11):e18614
URL: http://mhealth.jmir.org/2020/11/e18614/
doi:10.2196/18614
PMID:33146620

Original Paper

ReMindCare App for Early Psychosis: Pragmatic Real World Intervention and Usability Study

Lucia Bonet[1,2], MSci, PhD; John Torous[3], MD; David Arce[4], MSci; Ignacio Blanquer[4], Prof Dr; Julio Sanjuan[1,2,5], MD

[1]Department of Mental Health, Sanitary Research Institute of Valencia, University Clinic Hospital of Valencia, Valencia, Spain

[2]Faculty of Medicine and Odontology, University of Valencia, Valencia, Spain

[3]Division of Digital Psychiatry, Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

[4]Institute of Instrumentation for Molecular Imaging, Joint Centre of the Spanish National Research Council and Universitat Politècnica de València, Valencia, Spain

[5]Centre of Biomedical Investigation in Mental Health, Spanish Government Carlos III Health Institute, Madrid, Spain

Corresponding Author:
Julio Sanjuan, MD
Department of Mental Health
Sanitary Research Institute of Valencia
University Clinic Hospital of Valencia
Menéndez y Pelayo St, 4,
Valencia
Spain
Phone: 34 961 97 35 17
Email: julio.sanjuan@uv.es

Abstract

Background: eHealth interventions are widely used in clinical trials and increasingly in care settings as well; however, their efficacy in real-world contexts remains unknown. ReMindCare is a smartphone app that has been systematically implemented in a first episode of psychosis program (FEPP) for patients with early psychosis since 2018.

Objective: The objective of this study was to assess the efficacy of ReMindCare after 19 months of use in the clinic and varying use by individual patients.

Methods: The integration of the ReMindCare app into the FEPP started in October 2018. Patients with early psychosis self-selected to the app (ReMindCare group) or treatment as usual (TAU group). The outcome variables considered were adherence to the intervention and number of relapses, hospital admissions, and visits to urgent care units. Data from 90 patients with early psychosis were analyzed: 59 in the ReMindCare group and 31 in the TAU group. The mean age of the sample was 32.8 (SD 9.4) years, 73% (66/90) were males, 91% (83/90) were White, and 81% (74/90) were single.

Results: Significant differences between the ReMindCare and TAU groups were found in the number of relapses, hospitalizations, and visits to urgent care units, with each showing benefits for the app. Only 20% (12/59) of patients from the ReMindCare group had a relapse, while 58% (18/31) of the TAU patients had one or more relapses (χ^2=13.7, *P*=.001). Moreover, ReMindCare patients had fewer visits to urgent care units (χ^2=7.4, *P*=.006) and fewer hospitalizations than TAU patients (χ^2=4.6, *P*=.03). The mean of days using the app was 352.2 (SD 191.2; min/max: 18-594), and the mean of engagement was 84.5 (SD 16.04).

Conclusions: To our knowledge, this is the first eHealth intervention that has preliminarily proven its benefits in the real-world treatment of patients with early psychosis.

International Registered Report Identifier (IRRID): RR2-10.1111/eip.12960

(JMIR Mhealth Uhealth 2020;8(11):e22997) doi:10.2196/22997

KEYWORDS

app; clinical practice; mental health; psychosis; real-world intervention; telemedicine

Introduction

High interest in eHealth services and now digital and mobile health has been noted in many recent studies among patients with psychotic disorder diagnoses [1,2]. With COVID-19, this interest in digital health has surged, and the need to expand access to care through smartphones has become patent. Smartphone apps have been proposed as tools to mitigate social isolation, lack of access to care, and other triggers caused by the pandemic [3-5]. Researchers have already demonstrated that access to and use of technology among people with psychosis is nearly equivalent to that in the general population [6-8], but less is known about the actual efficacy of apps in care.

Apps have already seen growth in care for patients with early course psychosis. Many studies are using real-time ecological momentary assessment (EMA) surveys to monitor symptoms and experiences and identify early indicators of relapse [9]. Beyond relapse prediction, these EMA data can offer novel information on the longitudinal health status of patients, which could improve treatment and shared decision making between patient and physician [10]. Finally, eHealth services may be a major resource to enhance the benefits of the first episode of psychosis programs (FEPPs) for early psychosis, which can foster recovery [11] and reduce the risk of hospitalization and relapse [12,13].

Specific apps targeting schizophrenia have already been created and offer promising results. Examples of these innovative interventions are the Actissist [14] and the ExPRESS [15] interventions, which demonstrated potential in improving the quality of treatment of patients with early psychosis. Another example is the CrossCheck app [16], which demonstrated potential for identifying and dismantling dysfunctional beliefs that contribute to maintenance and distress associated with psychotic symptoms. Despite the widespread use of these eHealth interventions and high rates of efficacy reported in clinical trials, the efficiency and actual efficacy of these interventions in real-world clinical practice remains unknown [17].

One reason for the lack of initial success of health apps in clinical settings is lack of engagement. Often engagement in academic studies does not translate into real-world use [18,19]. Indeed, some studies found a negative correlation between the time spent using eHealth apps and the engagement of patients [20,21]. In addition, many clinicians expressed their concern that if these systems integrate seamlessly with clinical workflow, they will result in an increase in the clinicians' workload [22,23], which might affect their engagement with the app.

Other concerns have also limited efforts to integrate these apps into care settings. In our previous study [8], we found that 20% to 23% of patients felt anxious, suspicious, or paranoid concerning the internet, and almost 25% of patients perceived that use of the internet was directly related to one of their relapses. In addition, some studies indicated that excessive eHealth communications could be regarded as intrusive or irritating [24,25] or could increase worries about illness [25]. These potential harms of eHealth interventions must also be taken into consideration.

Considering these factors, it is clear that eHealth interventions shown to be feasible must now be assessed for effectiveness, efficacy, and efficiency [26] in real-world settings. With this objective in mind and to improve the daily treatment of patients with psychosis, we designed the ReMindCare app. The protocol followed for the design process and implementation of the app is published elsewhere [27]. In this protocol, we introduced ReMindCare as a smartphone app plus a clinician dashboard, developed to be implemented in a FEPP for patients with early psychosis.

To the best of our knowledge, ReMindCare is the first eHealth intervention for patients with early psychosis that has been systematically integrated into daily clinical practice, finally filling the gap between research and clinical practice [2,17].

The aim of this study was to assess the efficacy and clinical outcomes of the use of the app after 19 months in terms of adherence to ReMindCare, relapse prevention, hospital admissions, and visits to urgent care units compared with treatment as usual (TAU) without the app.

Methods

Study Setting

The app was systematically integrated into the daily clinical workflow in a FEPP at the University Clinic Hospital of Valencia, Spain. This FEPP started in 2010 with the objective of improving early detection, evaluation, and personalization of treatment. It covers a total of 330,000 inhabitants included in Area 5 of Valencia city. The incidence of novel psychotic disorders in this area has gradually increased during the 10 years since the program started. Currently, the FEPP in the clinic hospital has a mean of 30 to 35 new patients with psychosis per year.

The implementation of the ReMindCare app into the FEPP and into clinical practice started in October 2018 and is still in use today. In this study, we present the results from the first 19 months of use of the app.

Neither patients nor physicians received any remuneration or compensation for participating in the program or using the app. The use of the app was offered as an extra free service to the patients in the program.

Participants

Recruitment and Enrollment

The patient's psychiatrist of reference offered the use of the ReMindCare app to every outpatient from the FEPP who met the criteria for inclusion. Once patients enrolled in the study, they were encouraged to use the app as long as they remained in the program (maximum period of 5 years). To use the app, all patients signed an informed consent form and completed baseline assessments.

Eligibility Criteria

To be considered for this intervention, patients met the following criteria: (1) diagnosis of psychotic disorder following DSM-5 (*Diagnostic and Statistical Manual of Mental Disorders, 5th Edition*) criteria, interview conducted by a licensed clinician,

(2) aged between 17 and 65 years, (3) smartphone ownership with an internet connection that allows for the proper installation and functioning of the app, and (4) less than 5 years of illness duration. However, it must be stated that some patients remained in the program for more than 5 years. These patients remained in the FEPP to prevent potential relapses, as they experienced severe fluctuations in their symptoms.

Criteria for exclusion were (1) lack of ability to use and master a mobile device and the internet, (2) refusal to sign an informed consent form, and (3) level of Spanish or English not fluent enough to maintain a conversation or understand the app questionnaires.

Intervention

ReMindCare App

ReMindCare is a free and user-friendly app that conducts daily evaluations of the health status of patients with early psychosis by offering quick questionnaires (Figure 1).

Two types of questionnaires were included:

- Daily questionnaires: 3 daily questions assessing levels of anxiety, sadness, and irritability (Figure 2)
- Weekly questionnaires: 18 weekly questions aimed at assessing adherence to medication (1), the presence of side effects from antipsychotic medication intake (5), the attitude toward medication intake (3), and the presence of prodromal psychosis symptoms (9)

Figure 1. Screenshot of the ReMindCare app home screen.

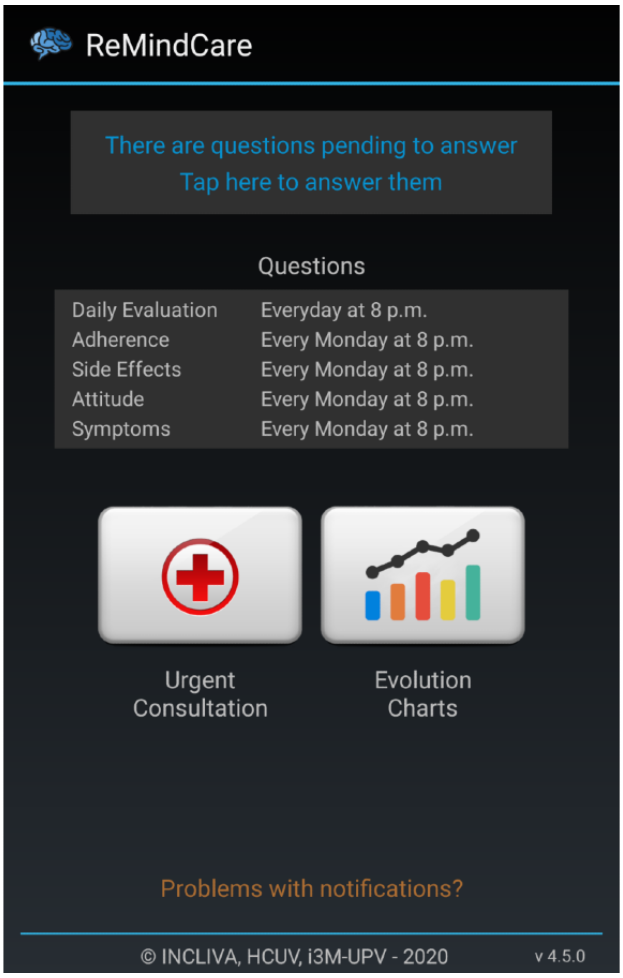
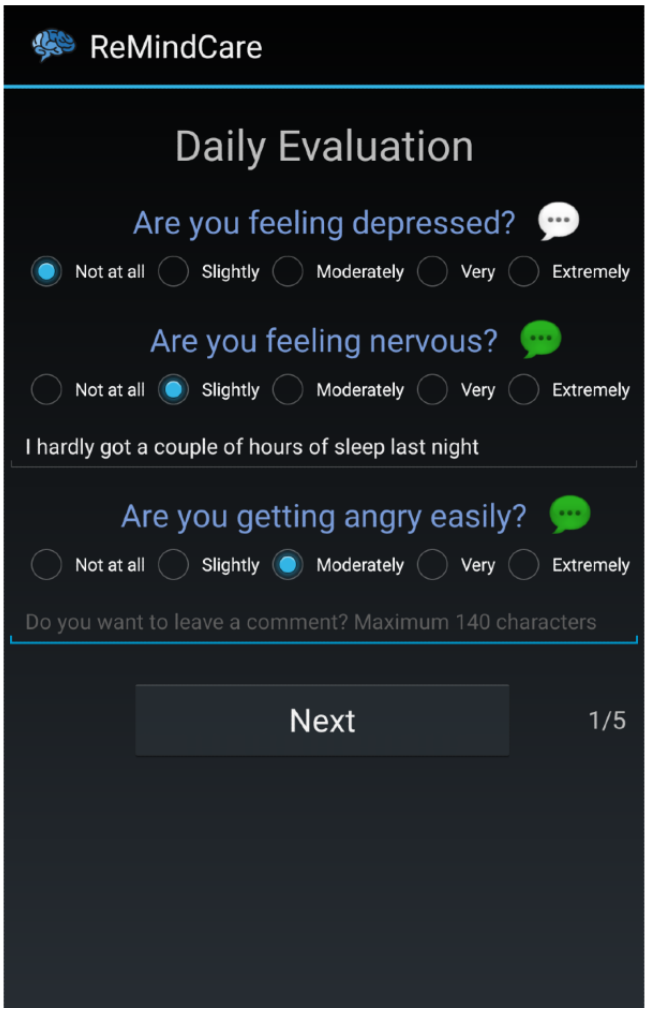


Figure 2. Screenshot of the ReMindCare daily questionnaire.



In addition, the app offered preset alerts in case of low engagement or abrupt changes in survey responses. Low engagement alerts were set off if patients did not respond to the surveys for 7 days or more, while abrupt changes were considered when there was a difference of 2 points (Likert scale 1 to 5) or more between each question in the last 2 surveys answered. These alerts notified physicians by email and were also displayed in the profile of the patient on the app's website portal.

All data captured by the app were accessible for physicians on a password-protected dashboard. Moreover, physicians could download a summary pdf of these data from the dashboard and attach it to the electronic clinical record of the patient in the hospital database.

The app is available in 3 languages (Spanish, English, and Catalán), although we are open to developing new language versions of the app. Our aim is to extend the use of the app to other countries, and adaptation of the app to different languages would be necessary to ensure patient engagement. Further information about the design process of the app and its characteristics can be found in the ReMindCare app study protocol [27].

Patients who used the app (ReMindCare group) did not experience any changes in their usual clinical appointments.

Treatment as Usual

The TAU group comprised patients who met the criteria but rejected using the app. In this group of patients, 42% (13/31) were patients with low adherence to treatment, 26% (8/31) did not perceive any benefit from using the app, and 26% (8/31) were suspicious about technology and their privacy. Additionally, 6% (2/31) were included in this group because they only used the app for 2 days. These patients continued with their usual psychiatric treatment at the FEPP and were not adversely affected by their rejection of participation.

Procedure

Once patients enrolled in the FEPP, after an interview with their psychiatrist of reference, they were asked to complete some baseline assessments. Subsequently, they were offered the use of the ReMindCare app. The ReMindCare app was described as an extra tool developed by the FEPP that could help them manage their symptoms and help clinicians better understand their illness evolution. The main characteristics of the app were listed. After receiving this information, patients decided whether they were willing to use the app. If they were not interested, they were placed in the TAU group. If patients were interested, they were informed in more detail by an expert clinician about the installation process, characteristics of the app, and ethics and data privacy information.

Patients could use the ReMindCare app to contact their psychiatrist of reference directly in case of symptoms worsening by using the urgent consultation request tab on the home screen of the app. If they clicked the urgent consultation request, their clinician would contact them by phone within 48 hours (patients who did not use the app could call the department of psychiatry at the hospital and be referred to their psychiatrist or attend an urgent care unit). In addition, clinicians contacted patients by phone in response to preset alarms. As a result of these phone calls and the information that patients provided to the clinician, urgent care visits could be scheduled if necessary. With these services, we aimed to improve the detection of early psychotic symptoms and reduce the visits to urgent care units at the hospital, as these prodromal symptoms will be primarily treated by a phone call or in the outpatient services. If patients did not make an urgent consultation request and no preset alarms were set off, they continued with their scheduled clinical appointments.

Furthermore, the use of the ReMindCare app changed the dynamics of the clinical appointment at the outpatient services. Once patients arrived at the clinical appointment, physicians accessed their profile on the ReMindCare's physician dashboard and used the information provided for patients to guide them through the interview. Clinicians used shared decision making with patients and discussed their responses.

Data Collection and Measures

Baseline

After patients were enrolled in the FEPP, the following data were collected:

- Sociodemographic information: age, gender, country, ethnicity, marital status, education level, employment status, and cohabitation
- Clinical information: antipsychotic medication, injectable medication, length of illness, associated illnesses, suicidal attempts, Clinical Global Impression Severity of Illness scale (CGI-SI) [28], Global Assessment of Functioning (GAF) [29], Positive and Negative Syndrome Scale (PANSS) [30], Premorbid Adjustment Scale (PAS) [31], date discharged from FEPP

Outcome Measures

- Efficacy: number of relapses, number of visits to the hospital urgent care units, and number of hospital admissions in the ReMindCare group compared with the TAU group
- Feasibility: number of patients who agreed to use the app compared with the patients who did not use it (TAU)
- Compliance and engagement: number of times patients answered the questionnaires when presented and number of months using the app, patients dropouts, plus number of urgent consultation requests

Data Analysis

Data were analyzed with the statistical program SPSS Statistics version 22 (IBM Corp). The cohort was divided into two groups: ReMindCare group patients agreed to use the app and used it for at least 1 month; the TAU group patients did not use the app or used it for less than 1 month. To consider that patients in the ReMindCare group had a relapse while using the app, patients had to be actively using the app. Relapses of patients who did not use the app for more than 2 months were not considered as relapses while using the app. Descriptive statistics (mean, standard deviation, frequency, and percentage) were determined, and chi-square test analysis was performed to compare the differences between the ReMindCare group and the TAU group.

Ethics, Data Privacy, and Participant Safety

The ReMindCare app project received approval from the research ethics committee of the faculty of medicine at the University of Valencia and from the research ethics committee of the Sanitary Research Institute of the University Clinic Hospital of Valencia, Spain.

To protect the data sent by patients, communications to the platform were encrypted with a transport layer security certificate from the Generalitat Valenciana and were sent through the https protocol. The hospital infrastructure is protected through a reverse proxy, which enhances security by establishing a single access point to it and hiding all inner infrastructures. Moreover, the integration of the app into the hospital systems was subjected to Organic Law 3/2018: protection of personal data and digital rights guarantee, December 5th, the Spanish organic law adaptation of the General Data Protection Regulation.

Results

Data from 90 patients were analyzed: 59 used or are using the app (ReMindCare group) and 31 did not agree to use the app (TAU group). Characteristics of both groups are displayed in Tables 1 and 2.

Table 1. Sociodemographic data.

Characteristic	Total	RC[a] group	TAU[b]	χ^2 (*P* value)
Age in years, mean (SD)	32.8 (9.4)	32.1 (1.2)	34.3 (1.7)	1.5 (.57)
24 and younger, n (%)	19 (21)	12 (20)	7 (23)	—[c]
25-44, n (%)	58 (64)	40 (68)	18 (58)	—
45 and older, n (%)	13 (14)	7 (12)	6 (19)	—
Gender (male), n (%)	66 (73)	40 (68)	25 (81)	1.7 (.19)
Native country (Spain), n (%)	79 (87)	48 (81)	30 (97)	4.2 (.04)
Race (White), n (%)	83 (91)	51 (86)	31 (100)	4.6 (.33)
Marital status, n (%)	—	—	—	5.2 (.16)
Single	74 (81)	50 (85)	23 (74)	—
Married	11 (12)	5 (9)	6 (19)	—
Other	85 (7)	4 (7)	2 (7)	—
Educational level, n (%)	—	—	—	5.9 (.05)
Primary	2 (2)	0 (0)	2 (7)	—
Secondary	45 (50)	27 (46)	18 (58)	—
College or higher	43 (48)	32 (54)	11 (36)	—
Employment status, n (%)	—	—	—	5.6 (.24)
Employed	29 (32)	16 (27)	13 (42)	—
Student	21 (23)	16 (27)	4 (13)	—
Not employed	38 (42)	25 (42)	13 (42)	—
Unable to work	3 (3)	2 (3)	1 (3)	—
Cohabitation, n (%)	—	—	—	2.3 (.51)
Alone	6 (7)	3 (5)	3 (10)	—
Family_birth	60 (66)	39 (66)	20 (65)	—
Family_own	11 (12)	6 (10)	5 (16)	—
Other	14 (15)	11 (19)	3 (10)	—

[a]RC: ReMindCare.

[b]TAU: treatment as usual.

[c]not applicable.

Table 2. Baseline clinical information.

Characteristics	Total	RC[a] group	TAU[b]	χ^2 (*P* value)
Injectable medication, n (%)	18 (20)	8 (14)	10 (32)	4.4 (.03)
Length of illness in years, mean (SD)	10.5 (2.8)	3.9 (0.4)	5.7 (0.5)	12.3 (.002)
0-1, n (%)	13 (14)	13 (22)	0 (0)	—[c]
2-5, n (%)	43 (48)	30 (51)	13 (42)	—
More than 6, n (%)	34 (38)	16 (27)	18 (58)	—
Associated illnesses, n (%)	29 (32)	18 (31)	11 (36)	0.2 (.63)
Suicidal attempts, n (%)	16 (18)	12 (22)	3 (10)	2.1 (.15)
CGI-SI[d], mean (SD)	4.2 (0.9)	4.1 (0.1)	4.4 (0.1)	2.7 (.26)
Mild (1-3), n (%)	13 (16)	10 (19)	3 (11)	—
Moderate (4-5), n (%)	66 (83)	42 (81)	24 (86)	—
Severe (>5), n (%)	1 (1)	0 (0)	1 (4)	—
GAF[e], mean (SD)	60.7 (10.9)	61.3 (1.7)	59.8 (1.7)	1.3 (.52)
Mild (71-100), n (%)	8 (10)	4 (8)	4 (14)	—
Moderate (51-70), n (%)	51 (65)	35 (69)	16 (57)	—
Severe (<50), n (%)	20 (25)	12 (24)	8 (29)	—
PANSS[f], mean (SD)	65.9 (18.8)	64.5 (2.2)	68.7 (4.6)	52.1 (.28)
Positive	18.4 (6.5)	18.7 (5.8)	18.7 (6.8)	23.9 (.58)
Negative	18.9 (6.9)	15.4 (5.1)	17.9 (9.3)	28.2 (.17)
N5. Difficulty in abstract thinking	2.3 (1.3)	2.0 (0.2)	2.8 (1.5)	12.8 (.03)
N6. Lack of spontaneity and flow conversation	1.7 (1.3)	1.6 (1.1)	1.9 (1.7)	12.9 (.02)
General	32.3 (8.2)	66.1 (14.7)	70.5 (22.2)	32.2 (.41)
G5. Mannerism and posturing	1.1 (0.7)	1.1 (0.4)	1.3 (0.7)	9.9 (.01)
PAS[g], mean (SD)	10.5 (2.8)	10.7 (0.5)	10.14 (0.6)	9.1 (.70)
Relapses_Baseline, n (%)	—	—	—	4.3 (.12)
0	53 (59)	38 (64)	15 (48)	—
1	21 (23)	14 (24)	7 (23)	—
≥2	16 (18)	7 (12)	9 (29)	—
UCU[h] visits_Baseline, n (%)	—	—	—	0.9 (.61)
0	26 (29)	19 (32)	7 (23)	—
1	36 (40)	23 (39)	13 (42)	—
≥2	28 (31)	17 (29)	11 (36)	—
Hospitalizations_Baseline, n (%)	—	—	—	4.6 (.10)
0	19 (21)	16 (27)	3 (10)	—
1	50 (56)	32 (54)	18 (58)	—
≥2	21 (23)	11 (19)	10 (32)	—

[a]RC: ReMindCare.

[b]TAU: treatment as usual.

[c]not applicable.

[d]CGI-SI: Clinical Global Impression Severity of Illness scale

[e]GAF: Global Assessment of Functioning.

[f]PANSS: Positive and Negative Syndrome Scale.

[g]PAS: Premorbid Adjustment Scale.

[h]UCU: urgent care units.

Sociodemographic Analysis

The mean age of the sample was 32.8 (SD 9.4) years, 73% (66/90) were males, 91% (83/90) were White, and 81% (74/90) were single. No significant differences were found between the ReMindCare and TAU groups in any of the sociodemographic information analyzed except for the native country. We found that nearly every immigrant considered for inclusion agreed to use the app (ReMindCare group 19% [10/11], TAU group 3% [1/11]; χ^2=4.2, *P*=.04). Further information regarding sociodemographic analysis of the data is displayed in Table 1.

Baseline Clinical Analysis

Significant differences were found between the ReMindCare group and TAU group in some clinical factors. With regard to injectable medication, 32% (10/31) of TAU patients were taking injectable medication, while only 14% (8/59) of the ReMindCare took it (χ^2=4.4, *P*=.04). Every new patient in the FEPP (length of illness: 0-1 year) agreed to use the app (13/90, 22%), and 58% (18/31) of the TAU group had their illness for 6 or more years (χ^2=12.3, *P*=.002). Moreover, the TAU patients showed higher scores on the PANSS N5 and N6 negative subscales and G5 in the general subscales (χ^2=12.8, *P*=.03; χ^2=12.9, *P*=.02; χ^2=9.9, *P*=.01, respectively).

Considering medication, 20% (18/90) of patients were taking injectable medications, 32% (29/90) of the patients suffered from another illness, and 18% (17/90) had a prior suicidal attempt. The mean of the CGI-SI was 4.2 (SD 0.9), the GAF mean=60.7 (SD 10.9), PANSS mean 65.9 (SD 18.8), and PAS mean 10.5 (SD 2.8). Finally, 12% (11/90) of patients were discharged from the FEPP. No significant differences were found between the groups in any of these factors. Moreover, no significant differences were found between the ReMindCare group and TAU group in terms of the number of relapses (χ^2=4.3, *P*=.12), visits to urgent care units (χ^2=0.9, *P*=.61), or the number of hospitalizations (χ^2=4.6, *P*=.10) at baseline. Further clinical information is available in Table 2.

ReMindCare Outcomes

The mean of days using the app was 352.2 (SD 191.2), which corresponds to 11.6 months. The mean of compliance was 84.5 (16.04), and 61.1% of the ReMindCare group had a compliance rate between 85% and 100%.

Of the 59 ReMindCare patients, 31% (18/59) requested an urgent consultation, 20% (12/59) had a relapse while using the app, and 8% (2/59) developed a delusion involving the app and the research group.

After 19 months of intervention, 63% (37/59) of patients continued using the app, while 12% (7/59) stopped using the app because they were discharged from the FEPP and 25% (15/59) opted to stop using ReMindCare. Reasons for discontinuation: 33% (5/15) of patients felt suspicious about technology (among these patients, 4 had a relapse while using the app); 40% (6/15) perceived the app as boring and did not perceive any benefit; and 27% (4/15) of patients left treatment and did not continue in the program. This information is shown in Table 3.

Table 3. Use of ReMindCare.

Characteristic	RC[a] group (n=59)	Min-max
Days using app, mean (SD)	352.2 (191.2)	18-594
Months using app, mean (SD)	11.6 (6.5)	0-19
Engagement, mean (SD)	84.5 (16.0)	42-100
85%-100%, n (%)	36 (61)	—[b]
UCU[c], n (%)	18 (31)	—
Relapses using app, n (%)	12 (20)	—
Relapses related to app, n (%)	2 (8)	—
Status of use after 19 months, n (%)		
Patients using app	37 (63)	—
Patients not using app	22 (37)	—
Discharged from FEPP[d]	7 (32)	—
Dropouts	15 (68)	—

[a]RC: ReMindCare.

[b]not applicable.

[c]UCU: urgent care units.

[d]FEPP: first episode of psychosis program.

With regard to the clinical outcomes, after 19 months of ReMindCare's integration into the clinical workflow, only 20% (12/59) of patients from the ReMindCare group had a relapse, while 58% (18/31) of TAU patients had one or more relapses

(χ^2=13.7, *P*=.001). Moreover, ReMindCare patients had fewer visits to urgent care units (χ^2=7.4, *P*=.006) and fewer hospitalizations than TAU patients (χ^2=4.6, *P*=.03). Information regarding these clinical outcomes is displayed in Table 4.

Table 4. Clinical outcomes after 19 months of the ReMindCare intervention.

Characteristic	Total, n (%)	RC[a] group, n (%)	TAU[b], n (%)	χ^2 (*P* value)
Relapses	—[c]	—	—	13.7 (.001)
0	60 (67)	47 (80)	13 (42)	—
1	29 (32)	12 (20)	17 (55)	—
≥2	1 (1)	0 (0)	1 (3)	—
UCU[d] visits	20 (22)	8 (14)	12 (39)	7.4 (.006)
Hospitalizations	9 (10)	3 (5)	6 (19)	4.6 (.03)

[a]RC: ReMindCare.

[b]TAU: treatment as usual.

[c]not applicable.

[d]UCU: urgent care units.

Discussion

Principal Findings

The results obtained from these analyses of the first 19 months of ReMindCare use highlight the potential benefits of this eHealth intervention for patients with early psychosis. Patients who used the app not only had fewer relapses than the TAU group, but they also had fewer visits to the urgent care unit and fewer hospitalizations.

Results related to the efficacy of the app are in line with previous results obtained in clinical trials [14-16]. However, as far as we know, this is the first study to identify the benefits of the use of an app as a tool systematically integrated into daily clinical practice in a FEPP.

With regard to the feasibility of the app, no significant differences were found between the ReMindCare group and the TAU group in terms of sociodemographic characteristics except for native country. The feasibility of this intervention aligns with the results obtained in our previous study [8], where we found no differences in terms of sociodemographic characteristics and interest in using eHealth interventions.

With regard to the clinical characteristics of the samples and their impact on the effect of ReMindCare, there were some differences between groups. We found that patients who did not use the app were more likely to be taking injectable medication, have a longer history of illness, and have higher scores on the PANSS N5 and N6 negative subscales and G5 in the general subscales. These results might suggest that the use of ReMindCare was not indicated for chronic patients. However, we did not find differences in other clinical scales such as the CGI-SI, GAF, and PAS scales or even on the PANSS total scale. More importantly, we did not find any differences between groups in terms of baseline relapses, hospitalizations, or visits to urgent care units.

These results are in line with the ones we obtained in our previous study [8], where we found that interest in using eHealth apps was equivalent between chronic and early psychosis patients. In this regard, we suggest that differences obtained in terms of the clinical characteristics of the patients could be more related to the history of treatment than to clinical characteristics. As we found, every new patient who joined the FEPP (length of illness less than 1 year) was interested in using the app (22% of users), while patients who had a longer history of treatment (length of illness more than 6 years) were more likely to reject its use (58% of TAU group). This could highlight the relevance of introducing these new technologies at the very beginning of treatment so early psychosis patients consider these apps to be just another tool included in their daily clinical treatment and not an extra service, especially since our results suggested that use of the app had a significant impact in improving the course of the illness.

Finally, with regard to compliance and engagement with the app, we found that 61% of patients had compliance rates between 85% to 100%. Rates of engagement were also high, as 63% of patients still use the app after almost 1 year. These results of compliance and long-term engagement are contrary to previous studies [20,21] and suggest that the use of an app in a long-term approach is feasible and beneficial.

However, we would like to highlight that 20% of patients had a relapse while using the app and 8% developed a delusion involving the use of the app and the research group. These negative results should be cautiously considered.

Technology could be a major resource to improve the quality of treatments, but as we found in a previous study [8], it can also play an important role as a trigger for psychotic symptoms. In this regard, in a 3-case study in 2011 conducted by Nitzan et al [32], they stated that the use of the internet and computers might contribute to a gradual break with reality and development of psychotic symptoms. They suggested that given that patients with psychotic diagnoses have greater difficulties in filtering and understanding signals and symbols, they are also more likely to misinterpret digital messages. However, no specific studies regarding the potential harms of the use of new technologies have been undertaken until the present.

In our study, we found that the ReMindCare app was related to beneficial clinical effects for the vast majority of patients who used it. However, despite the general positive effects found in this study, there are still some barriers and negative effects that must be taken into consideration. The main barrier found in our study relates to the 34% of the approached patients who did not want to use the app and who also tended to be the more chronic patients. Moreover, the main negative effect we found related to the 8% of patients who developed a delusion involving the app. As a result, we would like to point out that this app is not a panacea to prevent relapses. However, it is clear that the app positively affected the course of the illness, as only 5% of those who relapsed required hospitalization compared with 19% of patients who relapsed in the TAU group.

Limitations and Strengths

There were some limitations that must be taken into consideration. First, not every outpatient from the FEPP was eligible for inclusion, as some patients did not have their own smartphone with an internet connection or did not have the ability to use the app or understand it due to language barriers. Developing strategies to prevent digital exclusion should be a priority to ensure that every patient could benefit from these technologies [33]. Second, as a real-world study, this study was not randomized. Despite the groups not differing in the vast majority of clinical or demographic characteristics, there were some factors such as personality that could influence our results.

The main strength of our study was the fact that ReMindCare is the first app that has been systematically integrated into the clinical FEPP workflow. To our knowledge, there are no previous studies that used an app as a tool to improve the daily treatment of patients with early psychosis. All the studies we found were conducted in academic research settings that did not emulate real-world environments [17,34].

Another strength is in regard to the development of the ReMindCare app. First, it was based on two previous studies [2,8] and co-designed with patients [27]. Second, we conducted a pilot study and focus groups to ensure the involvement of both patients and care providers [27] in the design and improvement process of the app.

Finally, we would like to highlight the long-term approach of this intervention. As stated before, ReMindCare is now integrated into clinical practice and it was used for 19 months. These results align with previous studies [16] that found that people with psychosis have the abilities and interest required to engage in long-term eHealth interventions.

Implications for the Future

As a result of these analyses, we highlighted the benefits that the use of ReMindCare app produced on early psychosis patients in a FEPP. Our aim is to continue improving the app in response to the needs and suggestions provided by patients and clinicians. As Ross et al [22] claimed in their meta-review, in order to ensure the use of these eHealth technologies over time, there are three challenges that should be overcome. First, the apps must be able to adapt to the characteristics of the environment and patients. Second, the apps should be easy to use. Third, the apps should be integrated into clinical practice, adjusting the characteristics of the app in order to ensure it is user-friendly and efficient for patients and clinicians. It is our aim to address these issues to maintain the positive results obtained in this study.

However, we would like to point out a major issue that must guide future eHealth interventions. As stated before, 8% of patients developed a delusion related to the use of the app, 25% of patients deliberately stopped using the app, and 34% of patients approached did not want to use the app in the first place. These results suggest that there are still significant numbers of patients not willing to use eHealth interventions, and there are some patients who could be adversely affected by the use of these technologies. Studying the characteristics of these patients should guide future research in order to ensure that the use of digital technologies only provides benefits to the patients [8].

Finally, we would like to underline that given the exceptional situation that the world is facing at the moment with COVID-19 and in order to address the requirements of interventions that could improve the telematic treatment of patients and prevention of hospital collapses [4,35], ReMindCare could be used as an effective and efficient tool. Since quarantining in Spain began March 13, 2020, patients have not been permitted to come in person to their clinical appointments and have received their clinical evaluations by phone. Since that moment, the use of ReMindCare has been extremely useful to improve the evaluation and adherence of early psychosis patients. However, future analysis will be conducted in regard to this aspect.

As the conclusion of this study, we would like to point out that, to the best of our knowledge, ReMindCare is not only the first app to be integrated into the clinical practice, it is the first eHealth intervention with evidence that it improves the outcomes of early psychosis patients in a real-world care setting.

Acknowledgments

This study was supported by the Sanitary Research Institute of the University Clinic Hospital and the Mental Health Networking Biomedical Centre. It was also supported by the Generalitat Valenciana and the Program for Scientific Research, Technological Development, and Innovation in the Generalitat Valenciana of the European Union (2017/9830). It was also supported by grants from the Generalitat Valenciana (PROMETEO/2016/082, PROMETEO/2020/024), Carlos III Health Institute (PI13/00447; PI17/00402), and European Union through the European Regional Development Fund.

Authors' Contributions

LB and JS recruited and evaluated the patients. JS treated the patients, and LB supervised the app performance and the patients' responses and alarms. DA and IB developed the app and supervised its functioning, and they also drafted and discussed the paper. LB wrote the paper and analyzed the data. JS designed and supervised the project. JS and JT reviewed the data and supervised the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of "mHealth" among people with psychosis: a meta-analysis of cross-sectional studies. Schizophr Bull 2015 Sep 22;42(2):448-455. [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
2. Bonet L, Izquierdo C, Escartí MJ, Sancho JV, Arce D, Blanquer I, et al. Use of mobile technologies in patients with psychosis: a systematic review. Rev Psiquiatr Salud Ment 2017;10(3):168-178. [doi: 10.1016/j.rpsm.2017.01.003] [Medline: 28258835]
3. Kannarkat JT, Smith NN, McLeod-Bryant SA. Mobilization of telepsychiatry in response to COVID-19: moving toward 21 century access to care. Adm Policy Ment Health 2020 Jul;47(4):489-491 [FREE Full text] [doi: 10.1007/s10488-020-01044-z] [Medline: 32333227]
4. Torous J, Keshavan M. COVID-19, mobile health and serious mental illness. Schizophr Res 2020 Apr;218:36-37 [FREE Full text] [doi: 10.1016/j.schres.2020.04.013] [Medline: 32327314]
5. Wang C, Pan R, Wan X, Tan Y, Xu L, McIntyre R, et al. A longitudinal study on the mental health of general population during the COVID-19 epidemic in China. Brain Behav Immun 2020 Jul;87:40-48 [FREE Full text] [doi: 10.1016/j.bbi.2020.04.028] [Medline: 32298802]
6. Trefflich F, Kalckreuth S, Mergl R, Rummel-Kluge C. Psychiatric patients' internet use corresponds to the internet use of the general public. Psychiatry Res 2015 Mar 30;226(1):136-141. [doi: 10.1016/j.psychres.2014.12.037] [Medline: 25623020]
7. Gay K, Torous J, Joseph A, Pandya A, Duckworth K. Digital technology use among individuals with schizophrenia: results of an online survey. JMIR Ment Health 2016 May 04;3(2):e15 [FREE Full text] [doi: 10.2196/mental.5379] [Medline: 27146094]
8. Bonet L, Llácer B, Hernandez-Viadel M, Arce D, Blanquer I, Cañete C, et al. Differences in the use and opinions about new eHealth technologies among patients with psychosis: structured questionnaire. JMIR Ment Health 2018 Jul 25;5(3):e51 [FREE Full text] [doi: 10.2196/mental.9950] [Medline: 30045835]
9. Hau YS, Kim JK, Hur J, Chang MC. How about actively using telemedicine during the COVID-19 pandemic? J Med Syst 2020 Apr 30;44(6):108 [FREE Full text] [doi: 10.1007/s10916-020-01580-z] [Medline: 32350626]
10. Bucci S, Berry N, Morris R, Berry K, Haddock G, Lewis S, et al. "They are not hard-to-reach clients. We have just got hard-to-reach services." Staff views of digital health tools in specialist mental health services. Front Psychiatry 2019;10:344 [FREE Full text] [doi: 10.3389/fpsyt.2019.00344] [Medline: 31133906]
11. Arango C, Bernardo M, Bonet P, Cabrera A, Crespo-Facorro B, Cuesta MJ, et al. When the healthcare does not follow the evidence: the case of the lack of early intervention programs for psychosis in Spain. Rev Psiquiatr Salud Ment 2017;10(2):78-86. [doi: 10.1016/j.rpsm.2017.01.001] [Medline: 28189442]
12. Camacho E, Levin L, Torous J. Smartphone apps to support coordinated specialty care for prodromal and early course schizophrenia disorders: systematic review. J Med Internet Res 2019 Nov 12;21(11):e16393 [FREE Full text] [doi: 10.2196/16393] [Medline: 31714250]
13. Correll CU, Galling B, Pawar A, Krivko A, Bonetto C, Ruggeri M, et al. Comparison of early intervention services vs treatment as usual for early-phase psychosis: a systematic review, meta-analysis, and meta-regression. JAMA Psychiatry 2018 Jun 01;75(6):555-565 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.0623] [Medline: 29800949]
14. Bucci S, Barrowclough C, Ainsworth J, Machin M, Morris R, Berry K, et al. Actissist: proof-of-concept trial of a theory-driven digital intervention for psychosis. Schizophr Bull 2018 Mar 16;44(5):1070-1080. [doi: 10.1093/schbul/sby032] [Medline: 29566206]
15. Eisner E, Drake RJ, Berry N, Barrowclough C, Emsley R, Machin M, et al. Development and long-term acceptability of ExPRESS, a mobile phone app to monitor basic symptoms and early signs of psychosis relapse. JMIR Mhealth Uhealth 2019 Mar 29;7(3):e11568 [FREE Full text] [doi: 10.2196/11568] [Medline: 30924789]
16. Ben-Zeev D, Brian R, Wang R, Wang W, Campbell AT, Aung MSH, et al. CrossCheck: integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. Psychiatr Rehabil J 2017 Sep;40(3):266-275. [doi: 10.1037/prj0000243] [Medline: 28368138]
17. Torous J, Woodyatt J, Keshavan M, Tully LM. A new hope for early psychosis care: the evolving landscape of digital care tools. Br J Psychiatry 2019 May;214(5):269-272. [doi: 10.1192/bjp.2019.8] [Medline: 30739613]

18. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. J Affect Disord 2020 Feb;263:413-419. [doi: 10.1016/j.jad.2019.11.167]
19. Killikelly C, He Z, Reeder C, Wykes T. Improving adherence to web-based and mobile technologies for people with psychosis: systematic review of new potential predictors of adherence. JMIR Mhealth Uhealth 2017 Jul 20;5(7):e94 [FREE Full text] [doi: 10.2196/mhealth.7088] [Medline: 28729235]
20. Krzystanek M, Krysta K, Skałacka K. Treatment compliance in the long-term paranoid schizophrenia telemedicine study. J Technol Behav Sci 2017;2(2):84-87 [FREE Full text] [doi: 10.1007/s41347-017-0016-4] [Medline: 29082310]
21. Arnold C, Villagonzalo K, Meyer D, Farhall J, Foley F, Kyrios M, et al. Predicting engagement with an online psychosocial intervention for psychosis: exploring individual- and intervention-level predictors. Internet Interv 2019 Dec;18:100266 [FREE Full text] [doi: 10.1016/j.invent.2019.100266] [Medline: 31890619]
22. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
23. Allan S, Bradstreet S, Mcleod H, Farhall J, Lambrou M, Gleeson J, EMPOWER Group, et al. Developing a hypothetical implementation framework of expectations for monitoring early signs of psychosis relapse using a mobile app: qualitative study. J Med Internet Res 2019 Oct 24;21(10):e14366 [FREE Full text] [doi: 10.2196/14366] [Medline: 31651400]
24. Palmier-Claus JE, Rogers A, Ainsworth J, Machin M, Barrowclough C, Laverty L, et al. Integrating mobile-phone based assessment for psychosis into people's everyday lives and clinical care: a qualitative study. BMC Psychiatry 2013;13:34 [FREE Full text] [doi: 10.1186/1471-244X-13-34] [Medline: 23343329]
25. Kannisto KA, Adams CE, Koivunen M, Katajisto J, Välimäki M. Feedback on SMS reminders to encourage adherence among patients taking antipsychotic medication: a cross-sectional survey nested within a randomised trial. BMJ Open 2015;5(11):e008574. [doi: 10.1136/bmjopen-2015-008574] [Medline: 26553830]
26. Torous JB. Focusing on the future of mobile mental health and smartphone interventions. Psychiatr Serv 2018 Sep 01;69(9):945. [doi: 10.1176/appi.ps.201800308] [Medline: 30173627]
27. Bonet L, Torous J, Arce D, Blanquer I, Sanjuán J. ReMindCare, an app for daily clinical practice in patients with first episode psychosis: a pragmatic real-world study protocol. Early Interv Psychiatry 2020 Apr 06. [doi: 10.1111/eip.12960] [Medline: 32253830]
28. Busner J, Targum S. The clinical global impressions scale: applying a research tool in clinical practice. Psychiatry (Edgmont) 2007 Jul;4(7):28-37 [FREE Full text] [Medline: 20526405]
29. Endicott J, Spitzer RL, Fleiss JL, Cohen J. The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry 1976 Jun;33(6):766-771. [doi: 10.1001/archpsyc.1976.01770060086012] [Medline: 938196]
30. Peralta Martín V, Cuesta Zorita MJ. [Validation of positive and negative symptom scale (PANSS) in a sample of Spanish schizophrenic patients]. Actas Luso Esp Neurol Psiquiatr Cienc Afines 1994;22(4):171-177. [Medline: 7810373]
31. Cannon-Spoor HE, Potkin SG, Wyatt RJ. Measurement of premorbid adjustment in chronic schizophrenia. Schizophr Bull 1982;8(3):470-484. [doi: 10.1093/schbul/8.3.470] [Medline: 7134891]
32. Nitzan U, Shoshan E, Lev-Ran S, Fennig S. Internet-related psychosis−a sign of the times. Isr J Psychiatry Relat Sci 2011;48(3):207-211 [FREE Full text] [Medline: 22141146]
33. Greer B, Robotham D, Simblett S, Curtis H, Griffiths H, Wykes T. Digital exclusion among mental health service users: qualitative investigation. J Med Internet Res 2019 Jan 09;21(1):e11696 [FREE Full text] [doi: 10.2196/11696] [Medline: 30626564]
34. Rus-Calafell M, Schneider S. Are we there yet?! A literature review of recent digital technology advances for the treatment of early psychosis. Mhealth 2020;6:3 [FREE Full text] [doi: 10.21037/mhealth.2019.09.14] [Medline: 32190614]
35. Choi KR, Heilemann MV, Fauer A, Mead M. A second pandemic: mental health spillover from the novel coronavirus (COVID-19). J Am Psychiatr Nurses Assoc 2020;26(4):340-343. [doi: 10.1177/1078390320919803] [Medline: 32340586]

Abbreviations

CGI-SI: Clinical Global Impression Scale-Severity Illness
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
EMA: ecological momentary assessment
FEPP: first episode of psychosis program
GAF: Global Assessment of Functioning
PANSS: Positive and Negative Syndrome Scale
PAS: Premorbid Adjustment Scale
TAU: treatment as usual

Edited by L Buis; submitted 29.07.20; peer-reviewed by J Li, L Levin, MDG Pimentel; comments to author 15.09.20; revised version received 24.09.20; accepted 16.10.20; published 06.11.20.

Please cite as:
Bonet L, Torous J, Arce D, Blanquer I, Sanjuan J
ReMindCare App for Early Psychosis: Pragmatic Real World Intervention and Usability Study
JMIR Mhealth Uhealth 2020;8(11):e22997
URL: https://mhealth.jmir.org/2020/11/e22997
doi:10.2196/22997
PMID:33155986

Original Paper

Myofunctional Therapy App for Severe Apnea–Hypopnea Sleep Obstructive Syndrome: Pilot Randomized Controlled Trial

Carlos O'Connor-Reina[1,2*], MD, PhD; Jose Maria Ignacio Garcia[3,4*], MD, PhD; Elisa Rodriguez Ruiz[3*], BS, MS; Maria Del Carmen Morillo Dominguez[3*], MD; Victoria Ignacio Barrios[3,4*], MD; Peter Baptista Jardin[5*], MD, PhD; Juan Carlos Casado Morente[1,2*], MD, PhD; Maria Teresa Garcia Iriarte[6*], MD, PhD; Guillermo Plaza[7,8*], MD, PhD

[1]Otorhinolaryngology Department, Hospital Quironsalud Marbella, Marbella, Spain

[2]Otorhinolaryngology Department, Hospital Quironsalud Campo de Gibraltar, Palmones (Cadiz), Spain

[3]Pulmonology Department, Hospital Quironsalud Marbella, Marbella, Spain

[4]Pulmonology Department, Hospital Quironsalud Campo de Gibraltar, Palmones (Cadiz), Spain

[5]Otorhinolaryngology Department, Clinica Universitaria de Navarra, Pamplona, Spain

[6]Otorhinolaryngology Department, Hospital Universitario Virgen de Valme, Sevilla, Spain

[7]Otorhinolaryngogy Department, Hospital Universitario de Fuenlabrada, Universidad Rey Juan Carlos, Madrid, Spain

[8]Otorhinolaryngogy Department, Hospital Sanitas la Zarzuela, Madrid, Spain

[*]all authors contributed equally

Corresponding Author:
Carlos O'Connor-Reina, MD, PhD
Otorhinolaryngology Department
Hospital Quironsalud Marbella
Avda Severo Ochoa 22
Marbella
Spain
Phone: 34 952774200
Email: carlos.oconnor@quironsalud.es

Abstract

Background: Myofunctional therapy has demonstrated efficacy in treating sleep-disordered breathing. We assessed the clinical use of a new mobile health (mHealth) app that uses a smartphone to teach patients with severe obstructive sleep apnea–hypopnea syndrome (OSAHS) to perform oropharyngeal exercises.

Objective: We conducted a pilot randomized trial to evaluate the effects of the app in patients with severe OSAHS.

Methods: Forty patients with severe OSAHS (apnea–hypoxia index [AHI]>30) were enrolled prospectively and randomized into an intervention group that used the app for 90 sessions or a control group. Anthropometric measures, Epworth Sleepiness Scale (0-24), Pittsburgh Sleep Quality Index (0-21), Iowa Oral Performance Instrument (IOPI) scores, and oxygen desaturation index were measured before and after the intervention.

Results: After the intervention, 28 patients remained. No significant changes were observed in the control group; however, the intervention group showed significant improvements in most metrics. AHI decreased by 53.4% from 44.7 (range 33.8-55.6) to 20.88 (14.02-27.7) events/hour (P<.001). The oxygen desaturation index decreased by 46.5% from 36.31 (27.19-43.43) to 19.4 (12.9-25.98) events/hour (P=.003). The IOPI maximum tongue score increased from 39.83 (35.32-45.2) to 59.06 (54.74-64.00) kPa (P<.001), and the IOPI maximum lip score increased from 27.89 (24.16-32.47) to 44.11 (39.5-48.8) kPa (P<.001). The AHI correlated significantly with IOPI tongue and lip improvements (Pearson correlation coefficient −0.56 and −0.46, respectively; both P<.001). The Epworth Sleepiness Scale score decreased from 10.33 (8.71-12.24) to 5.37 (3.45-7.28) in the app group (P<.001), but the Pittsburgh Sleep Quality Index did not change significantly.

Conclusions: Orofacial exercises performed using an mHealth app reduced OSAHS severity and symptoms, and represent a promising treatment for OSAHS.

Trial Registration: Spanish Registry of Clinical Studies AWGAPN-2019-01, ClinicalTrials.gov NCT04438785; https://clinicaltrials.gov/ct2/show/NCT04438785

(JMIR Mhealth Uhealth 2020;8(11):e23123) doi:10.2196/23123

KEYWORDS

myofunctional therapy; oropharyngeal exercises; mHealth; sleep apnea; smartphone app; app; sleep; therapy; apnea; randomized trial; efficacy

Introduction

Background

Obstructive sleep apnea–hypopnea syndrome (OSAHS) is a serious health problem worldwide [1], and is associated with morbidities such as hypertension, arrhythmia, and cerebrovascular diseases. The classic treatment of this syndrome is based on dietary measures, weight loss, and exercise, and the use of continuous positive airway pressure (CPAP). Other treatment options include upper airway surgery, mandibular advancement devices, and upper airway stimulation devices (UASDs) that bring the tongue forward to prevent it from falling backward and collapsing the airway. The success rates of treating the airway obstacle or correcting the muscles vary. Indications and the success rate for all treatments depend on patient compliance with the treatment and the severity of the disease [2].

Patients with OSAHS present with impaired sensorimotor deficits located in the upper airway muscles [3]. These deficits are associated with apraxia [4], hypotonia [5], and changes in the type of muscle fibers from type I to types IIa and IIb, which lead to early fatigue and can interfere with the ability to perform prolonged exercise [6]. These deficits can lead to impairment of proprioceptive acuity in the upper airway muscles [4]. The best rehabilitation for improving this pathology is proprioceptive training in association with visual or acoustic feedback [7].

Myofunctional therapy is one of the newest treatments for sleep-disordered breathing [8], which is based on daily exercises using the oropharyngeal muscles in an attempt to strengthen them and to facilitate opening of the airway. OSAHS originates from suboptimal function of the dilator muscles of the airway. Therefore, myofunctional therapy is designed, theoretically, to deal with the underlying mechanism of this disease [9]. The patient is instructed to perform these exercises regularly for 20-40 minutes daily for at least 3 months [10]. In some cases, patients perform the exercises independently at home without substantial feedback and without giving precise information to the therapist about their performance of the exercises, and this type of therapy is associated with low adherence [11].

Most existing mobile health (mHealth) apps for OSAHS focus on the diagnosis of snoring or OSAHS [12], and a few are designed to promote adherence to treatment with CPAP [13]. None of these apps is used alone to treat OSAHS. Mobile technology may be valuable for treating people with OSAHS because it may promote patient empowerment and self-management [12].

Therefore, we conducted this pilot randomized trial to evaluate a new mHealth app based on proprioceptive training, which was designed to promote oropharyngeal exercises through interactions with a smartphone. In this prospective, randomized, multicenter clinical study, we evaluated adherence to the app and its effectiveness in a group of patients with severe OSAHS, as identified by an apnea–hypoxia index (AHI)>30 compared with a control group of similar patients who did not participate in the intervention.

Objectives

Primary Objectives

The main objectives were to study the effects of the AirwayGym app on adherence to myofunctional therapy and on the AHI in patients recently diagnosed with severe OSAHS (AHI>30).

Secondary Objectives

The secondary objectives were to evaluate the change in the oxygen desaturation index (ODI), use of the Iowa Oral Performance Instrument (IOPI) score to evaluate the effects of the app on the tone of the genioglossus and buccinator muscles, and use of the Epworth Sleepiness Scale and Pittsburgh Sleep Quality Index questionnaires to evaluate subjective morning somnolence and sleep quality.

Methods

Study Overview

This was a nonsponsored study coordinated by the Pulmonology and Otolaryngology Departments of Quirónsalud Marbella Hospital and Campo de Gibraltar Hospital, Andalucia, Spain. The protocol was designed and written by the authors and is available in Multimedia Appendix 1. The protocol was approved by the governmental review board (AWGAPN-2019-01). All patients provided informed consent. An independent data and safety monitoring board regularly reviewed data on serious or nonserious adverse events and study quality. The authors vouch for the accuracy and completeness of the data reported and for the fidelity of the study to the protocol. All investigators were GCP-certified.

The CONSORT (Consolidated Standards of Reporting Trials) checklist [14] is shown in Multimedia Appendix 2.

Design

This was a prospective controlled quasiexperimental clinical study in patients with severe OSAHS (AHI>30).

Participants and Recruitment

Patients newly diagnosed with severe OSAHS based on the results of polysomnography or respiratory polygraphy with measures of AHI and oxygen saturation were recruited offline in a clinical setting. All sleep studies were interpreted manually by a sleep technician according to the standard criteria of the American Academy of Sleep Medicine Manual for the Scoring of Sleep and Associated Events [15], and the interpretations were reviewed by certified physicians.

Information about the inclusion and exclusion criteria, evaluation of the type of smartphone used, previous experience with the app, and the study protocol are provided in Multimedia Appendix 1.

All patients agreed to participate and provided offline informed consent. At the initial visit, participants were evaluated by an otorhinolaryngologist who performed rhinofibrolaryngoscopy, Friedman staging, the Marchesani protocol [16], and examination of temporomandibular joint dysfunction. Patients with grade IV tonsils, complete nasal obstruction, ankyloglossia, or problems with temporomandibular joint dysfunction were excluded from the study. In a second visit, anthropometric variables, including weight, height, and neck and waist circumferences, were measured, and the BMI was calculated. The Friedman staging, Epworth Sleepiness Scale, and Pittsburgh Sleep Quality Index questionnaires were completed, and IOPI lingual and buccinator scores were obtained.

Randomization

Randomization was based on the consecutive order of patient enrollment. A pulmonologist specialist allocated odd-numbered patients to the AirwayGym app group and even-numbered patients to the control group.

Intervention

Participants in the AirwayGym group were instructed about the use of the app and the exercises to perform for 20 minutes daily. Follow-up visits for both the AirwayGym and control groups occurred after 1 month (visit 3) and 3 months (visit 4). At these visits, all variables were measured again and the questionnaires were completed, and the patients were asked whether they were using any other therapies. In the final visit at 3 months, polysomnography or polygraphy was performed for both groups. The total study duration for each participant was 3 months.

Participants were excluded from the study if they were lost to follow up because they did not attend hospital visits or if they lost ≥5% of body weight during the study. Patients in the AirwayGym group were also excluded if they did not perform ≤85% of the scheduled exercise sessions, as monitored by the app.

Myofunctional Therapy App

Participants in the AirwayGym group committed to using the AirwayGym app. This app was created as a collaboration between the sleep units of Quirónsalud Marbella Hospital and Campo de Gibraltar Hospital. The app can be thought of as a portable fitness app except that the user is intended for patients rather than athletes, and therapists rather than trainers provide the instructions. The novelty of this app is that it is the first app in the health care market that allows the patient to interact directly with the smartphone without needing any other device. The app focuses on sleep apnea disease and improving proprioceptive deficits. When used with the app, the phone provides acoustic feedback about the efficacy of the exercises performed.

The app includes 9 exercises based on myofunctional therapy (Figure 1) that are aimed at improving the tonicity of the various muscles involved in the pathogenesis of OSAHS. Before every exercise, an animated demonstration and a video show the patient how to perform the exercise. After each exercise, the patient receives visual, acoustic, and tactile feedback about the success of their performance as a point score. When the patient finishes the exercises, the results are saved on a networked online storage (in the cloud), and a therapist can evaluate the patient's performance of the exercises. Users of the app can follow the progress of their daily activity over time (Figure 2). A chat function is available through which the patient can contact the therapist directly. Additional information can be found on the AirwayGym webpage [17]. This app complies with regulations 2002/58/CE and (UE) 2016/679 concerning data protection. The app was provided free to each participant.

The main objective of the exercises in the app is to increase the tone of the extrinsic muscles of the tongue (genioglossus, hyoglossus, styloglossus, and palatoglossus). The exercises are based on those described elsewhere [10] and have been adapted to allow feedback using a smartphone (see Multimedia Appendices 3-12).

Figure 1. Screenshot of a gif showing an exercise.

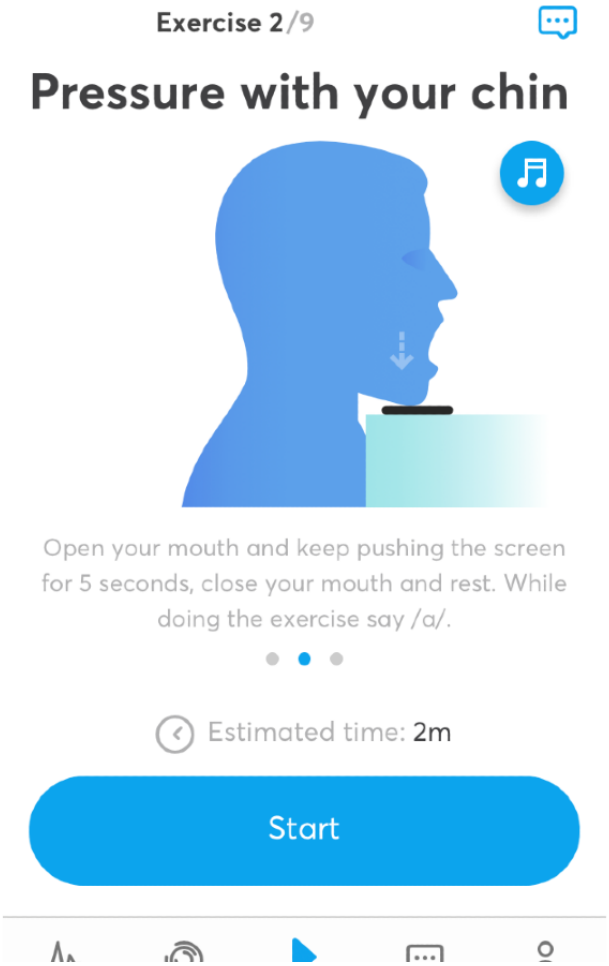
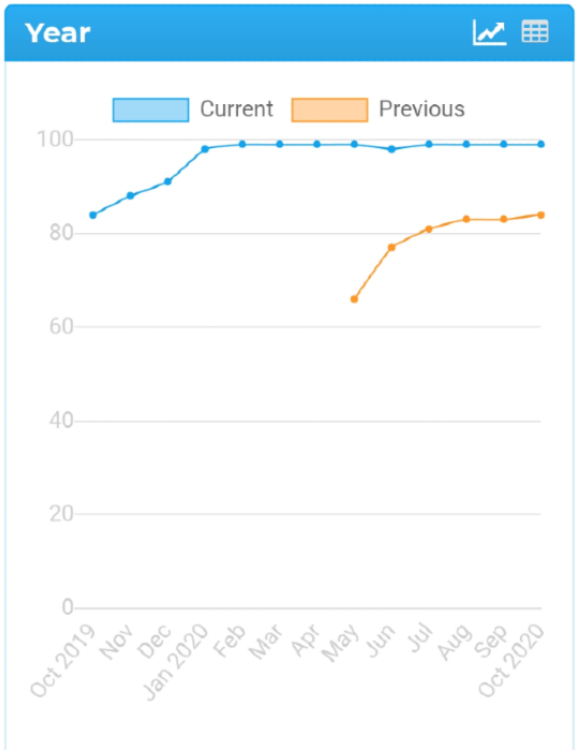


Figure 2. Screenshot of a patient's progress in following the exercises for 1 year.



Sleep Laboratory Procedures

Standard laboratory polysomnography (Somté PSG, Compumedics Ltd 2006, Abbotsford, Australia) was performed according to the technical specifications of the American Academy of Sleep Medicine [15]. The recorded variables were obtained using electroencephalography (C3-A2, C4-A1, O1-A2, O2-A1), electrooculography (2 channels), chin and leg electromyography, and electrocardiography. Frontal electrodes were not used. Respiratory variables were measured using linearized nasal pressure prongs and the flow waveform of the oronasal thermal signals. Respiratory effort signals were measured through inductive bands that recorded ribcage and abdominal movements. Oxygen saturation, body position, and snoring were also registered.

Respiratory polygraphy was performed using an Embletta portable diagnostic system (ResMed, Sydney, Australia) according to the technical specifications of the American Academy of Sleep Medicine [15]. Measurements were obtained using a snoring sensor, nasal thermistor, and nasal pressure cannula to register airflow; thoracic and abdominal belts to assess ribcage and abdominal movements; electrocardiography; actigraphy to detect body position; oxygen saturation; and heart rate.

All patients were evaluated using the same testing procedure (polysomnography or respiratory polygraphy) before and after the intervention. The results for each participant were analyzed manually by a technician who was blinded to the participant's assigned group.

The pulmonologist's medical evaluation was used to determine which test was chosen for each patient. Apnea and hypopnea were analyzed and scored according to the following criteria. Hypopnea was defined as a ≥30% decrease in airflow signal amplitude lasting ≥10 seconds and accompanied by ≥3% oxygen desaturation. Apnea was defined as a ≥90% decrease in airflow signal amplitude lasting ≥10 seconds. The ODI was used to quantify oxygen desaturation ≥3% Both tests were used to define moderate OSAHS as an AHI of 15-29.9 events/hour of sleep and severe OSA as ≥30 events/hour of sleep.

IOPI Measurements

Detailed information about this device and measurements is provided in Multimedia Appendix 1.

Sample Size Derivation

The effectiveness of use of the app for performing myofunctional therapy in patients with severe OSAHS was evaluated using the percentage changes in the AHI observed during follow up as the primary outcome measure. This percentage was calculated from results reported in previous studies of myofunctional therapy [10,18,19]. Based on an α level of .05 and power of 0.80, we estimated that 30 participants were required. To account for potential loss during the inclusion process (including patients with selection bias), early withdrawal, or loss to follow up, we doubled the sample size to 60. The sample size was calculated using XLSTAT (v16 Addinsoft France).

Data Analysis

Data were collected in a database. Nominal variables are described by their frequency distribution. Quantitative variables were assessed by calculating the median and IQR. Baseline characteristics of the 2 groups of patients with OSAHS were compared using two-tailed paired *t*-tests for continuous variables and the chi-squared or Fisher exact test for nominal variables. For variables with a skewed distribution, the Mann-Whitney *U* test was used. Pearson correlational analysis was used to assess the associations between the changes in the AHI and changes in possible explanatory variables, including BMI and neck and waist circumferences. A *P* value <.05 was considered to be significant.

Results

Patients

Of the 60 patients initially recruited, 40 patients were enrolled and randomized from February 2019 to July 2020. Twenty of the 60 patients were excluded, 10 (17%) because of the exclusion criteria and 10 (17%) because of findings in the otorhinolaryngologist's examination. Six of the 40 patients were excluded because of a change in body weight, 4 voluntarily abandoned the study in the control group, and 2 patients were lost because of poor adherence to therapy in the AirwayGym group. Finally, the data for 28 patients (22 men) were included in the study (Figure 3). Half of the participants in the control group were lost to follow up.

The baseline demographic and sleep characteristics are presented for the two groups in Table 1. There were no significant differences in any characteristics, including age, weight, and BMI, between the two groups. In addition, at the baseline, the AHI did not differ within or between the groups (Table 1).

Figure 3. CONSORT flowchart for pilot randomized trials for recruitment of participants in this study. AGG: AirwayGym group; CG: control group.

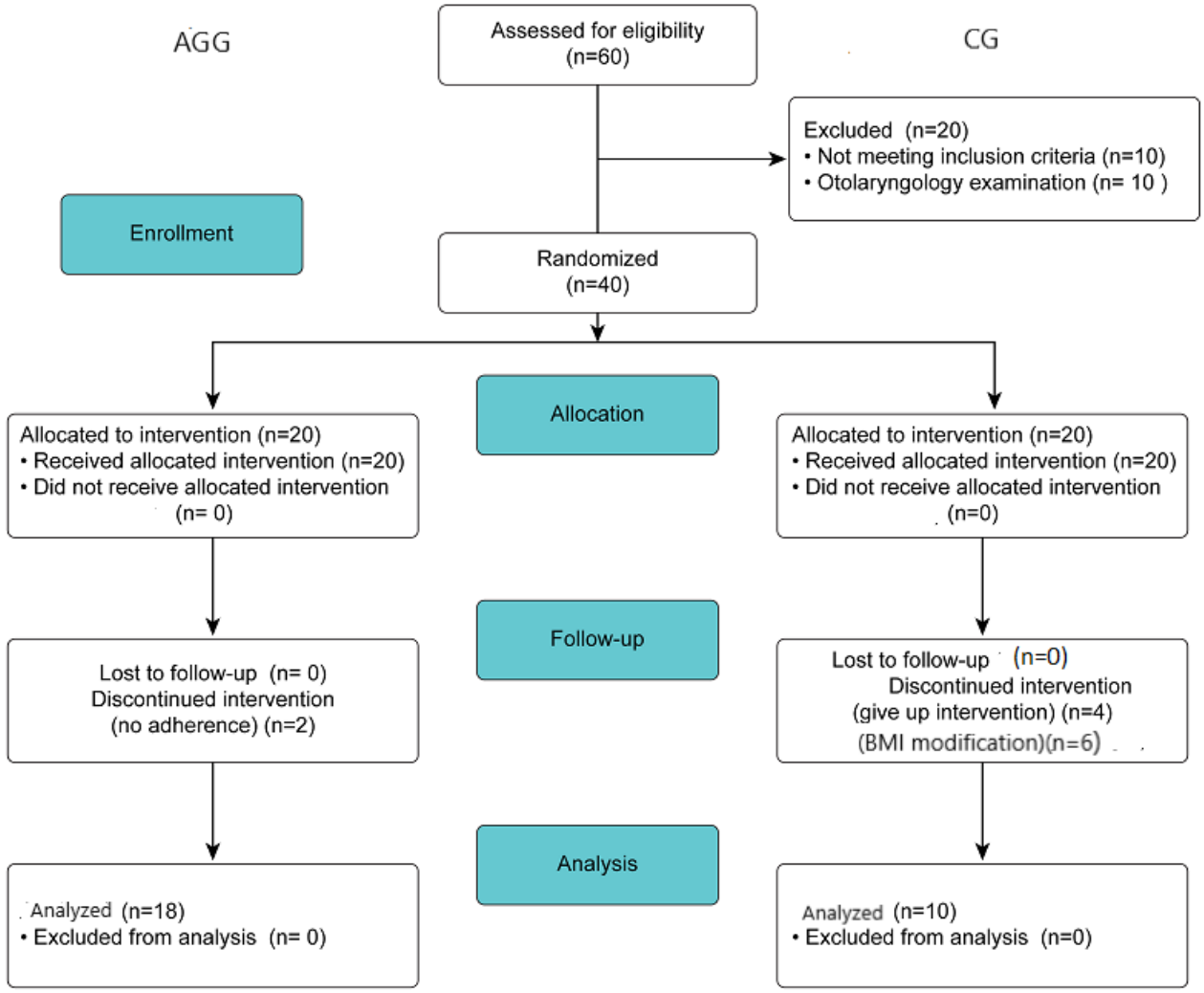


Table 1. Baseline characteristics of the study participants.

Characteristic	Control (n=10)	AirwayGym (n=18)	*P* value
Anthropometric data			
Age (years), median (IQR)	63.9 (56.4-71.38)	59.17 (53.7-64.6)	.19
Female, n (%)	2 (20)	4 (22)	.95
Body weight (kg), median (IQR)	87.7 (77.9-97.4)	86.11 (77.4-94.8)	.86
Height (cm), median (IQR)	170.2 (166.2-174.1)	177.1 (166.3-177.9)	.52
Waist circumference (cm), median (IQR)	109.1 (97.5-120.6)	109.2 (102.9-115.4)	.98
Neck circumference (cm), median (IQR)	43.7 (40.3-47)	43.74 (41.03-46.4)	.99
BMI (kg/m^2), median (IQR)	29.6 (27.1-32.08)	28.9 (26.8-31.02)	.67
Friedman stage, n (stage)	2 (I), 3 (II), 2 (III), 3 (IV)	4 (I), 4 (II), 3 (III), 7 (IV)	.95
Polysomnography data			
AHI[a]/hour, median (IQR)	47.36 (38.59-56.13)	44.77 (33.841-55.69)	.70
ODI[b], median (IQR)	40.6 (29.46-51.81)	36.31 (27.1-43.43)	.54
Tone measures (kPa)			
IOPI[c] max tongue[d], median (IQR)	42 (32.67-51.33)	40.26 (35.32-45.2)	.70
IOPI max lips[e], median (IQR)	28.10 (23.7-32.5)	28.3 (24.16-32.47)	.94
Questionnaires			
Pittsburgh Sleep Quality Index, median (IQR)	8.80 (7.12-10.48)	10.1 (8.22-11.99)	.34
Epworth Sleepiness Scale, median (IQR)	9.3 (6.6-12)	10.47 (8.71-12.24)	.42

[a]AHI: apnea–hypopnea index.

[b]ODI: oxygen desaturation index.

[c]IOPI: Iowa Oral Performance Instrument.

[d]maximum tongue elevation strength.

[e]maximum lip strength.

Intervention

After the intervention period, none of the variables changed significantly in the control group (Table 2). Conversely, significant changes from before to after the intervention were observed in the AirwayGym group. The adherence to the therapy in the AirwayGym group was 90%. No adverse reactions were registered.

Table 2. Changes in variables from the baseline to 3-month follow up in the control and AirwayGym groups.

Variable	Control group (n=10)			AirwayGym group (n=18)		
	Baseline, median (IQR)	After 3 months, median (IQR)	*P* value	Baseline, median (IQR)	After 3 months, median (IQR)	*P* value
Anthropometric data						
Body weight (kg)	87.7 (77.9-97.4)	87.3 (78.03-95.66)	.95	86.1 (77.4-94.8)	86.00 (77.93-94.06)	.92
Height (cm)	170.2 (166.2-174.1)	169.80 (166.27-174.12)	.98	177.1 (166.3-177.9)	172.6 (167.3-177.9)	1.00
Waist circumference (cm)	109.1 (97.5-120.6)	108.5 (96.9-120.10)	.94	109.2 (102.9-115.4)	108.84 (102.7-114.9	.92
Neck circumference (cm)	43.7 (40.3-47)	43.5 (40.33-46.67)	.92	43.74 (41.03-46.4)	44.6 (42.12-47.2)	.60
BMI (kg/m^2)	29.63 (27.1-32.08)	29.6 (27.35-31.84)	.94	28.9 (26.8-31.02)	28.81 (26.79-30.83)	.92
Polysomnography data						
AHI[a] (events/hour)	47.36 (38.59-56.13)	35.00 (31.2-38.7)	.07	44.77 (33.84-55.69)	20.88 (14.02-27.74)	<.001
ODI[b]	40.64 (29.46-51.81)	32.03 (24.14-39.91)	.17	36.31 (27.19-43.43)	19.4 (12.9-25.98)	.003
Tone measures (kPa)						
IOPI[c] max tongue[d]	42 (32.67-51.33)	44.2 (34.1-54.2)	.72	39.83 (35.32-45.2)	59.06 (54.74-64.00)	<.001
IOPI max lips[e]	28.10 (23.7-32.5)	31.3 (26.6-35.9)	.27	27.89 (24.16-32.47)	44.11 (39.5-48.8)	<.001
Questionnaires						
Pittsburgh Sleep Quality Index	8.80 (7.12-10.48)	9.78 (7.2-12.11)	.50	10.2 (8.22-11.99)	8.28 (5.97-10.35)	.22
Epworth Sleepiness Scale	9.3 (6.6-12.0)	9.6 (6.8-12.3)	.86	10.33 (8.71-12.24)	5.37 (3.45-7.28)	<.001

[a]AHI: apnea–hypopnea index.

[b]ODI: oxygen desaturation index.

[c]IOPI: Iowa Oral Performance Instrument.

[d]maximum tongue elevation strength.

[e]maximum lip strength.

Anthropometric Data

The anthropometric measures did not change significantly in the intervention group (Table 2).

Laboratory and IOPI Measurements

The AHI decreased by 53.36% from 44.7 (IQR 33.8-55.6) to 20.88 (IQR 14.02-27.7) events/hour (*P*<.001) (Figure 4). The ODI score decreased by 46.5% from 36.31 (IQR 27.19-43.43) to 19.4 (IQR 12.9-25.98) events/hour (*P*=.003). The IOPI maximum tongue elevation strength score increased from 39.83 (IQR 35.32-45.20) to 59.06 (IQR 54.74-64.00) kPa (*P*<.001) (Figure 5). The IOPI maximum lip score increased from 27.89 (IQR 24.16-32.47) to 44.11(IQR 39.5-48.80) kPa (*P*<.001) (Figure 6). The AHI correlated significantly with improvements in the IOPI tongue and lip scores (Pearson coefficients −0.56, *P*<.001 and −0.46, *P*<.001, respectively).

Figure 4. Apnea–hypopnea index (AHI) before (baseline) and after 3 months in patients with severe obstructive sleep apnea. Intragroup comparison from before to after the study was performed using the Wilcoxon test owing to the skewed data distribution. CG: control group; AGG: AirwayGym group.

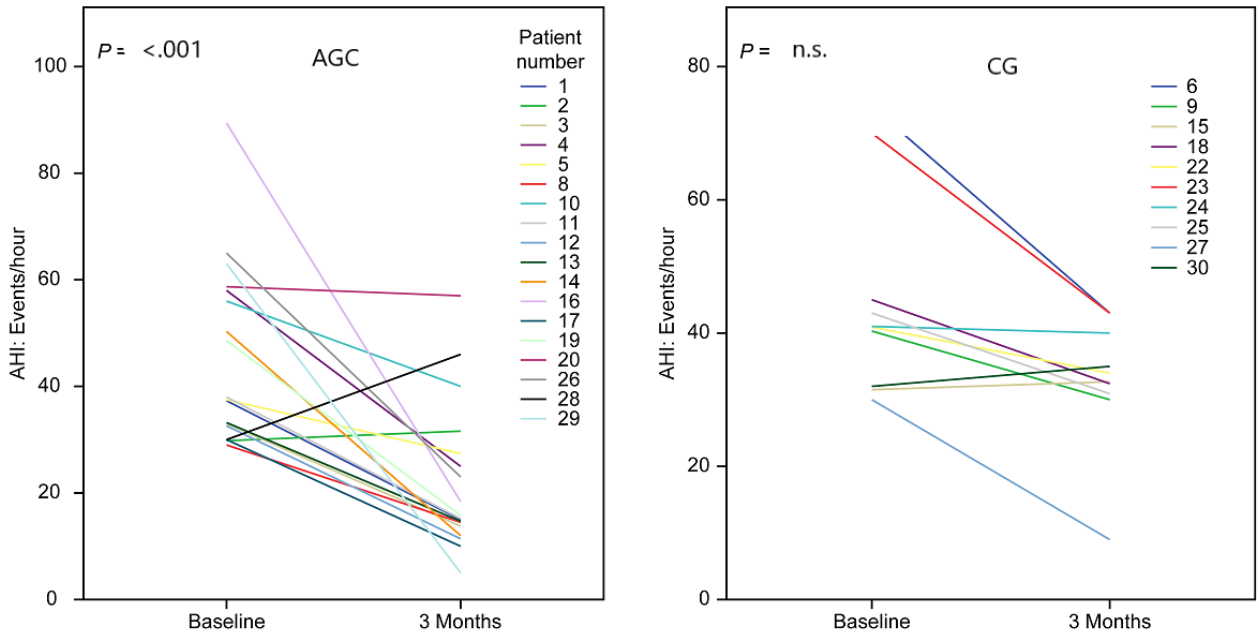


Figure 5. Iowa Oral Performance Instrument (IOPI) tongue strength (kPa) at baseline and after 3 months in patients with severe obstructive sleep apnea. The intragroup comparison is shown from before to after the study. AGG: AirwayGym group; CG: control group.

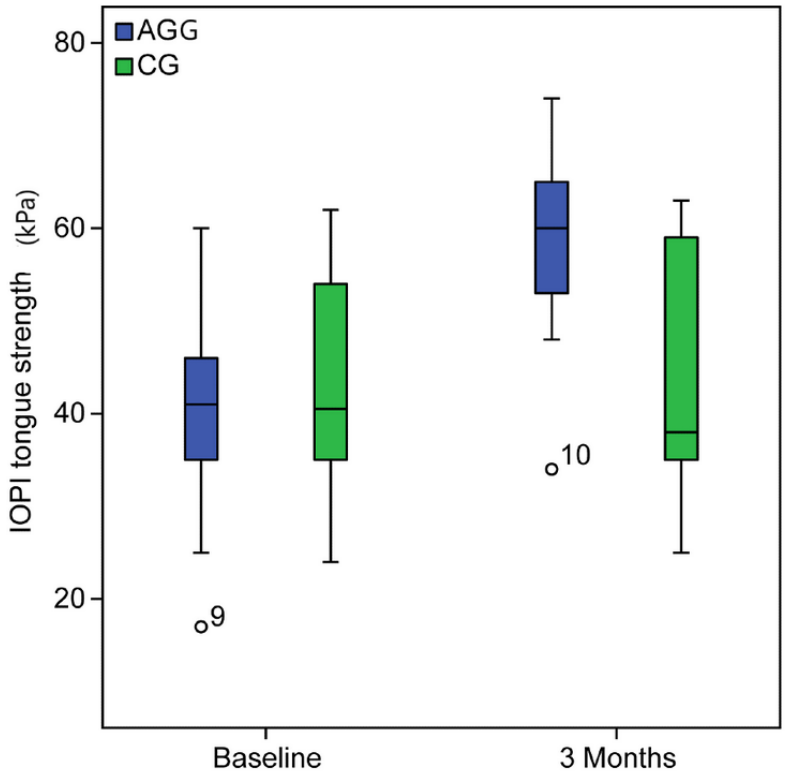
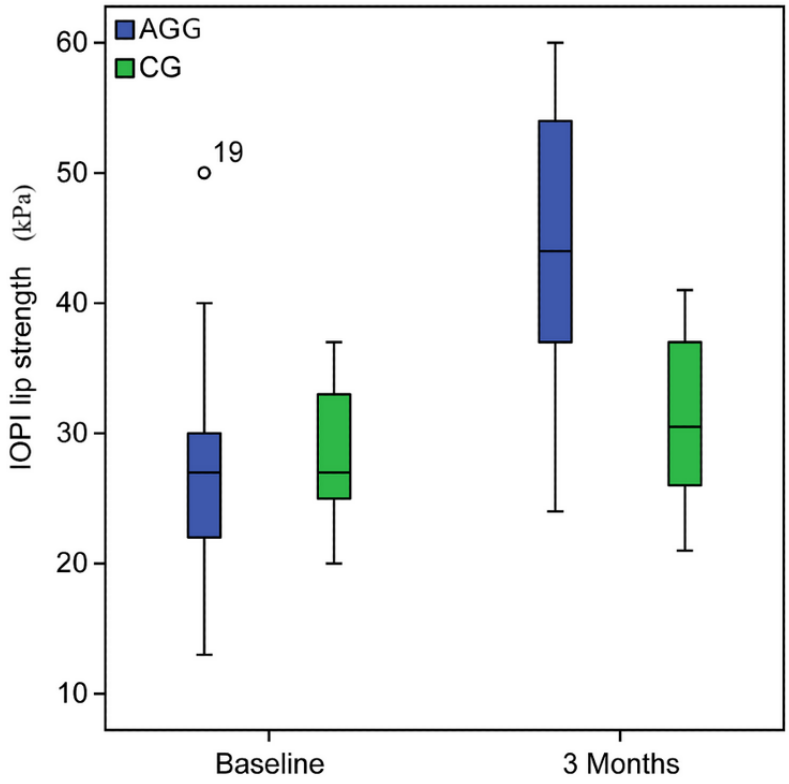


Figure 6. Iowa Oral Performance Instrument (IOPI) lip strength (kPa) at baseline and after 3 months in patients with severe obstructive sleep apnea. The intragroup comparison is shown from before to after the study. AGG: AirwayGym group; CG: control group.



Questionnaires

The Epworth Sleepiness Scale score decreased from 10.33 (IQR 8.71-12.24) to 5.37 (IQR 3.45-7.28) (P<.001). However, the Pittsburgh Sleep Quality Index did not change significantly in the AirwayGym group.

Discussion

Principal Findings

This prospective study showed that an mHealth app that includes education about myofunctional therapy exercises for patients with severe OSAHS helped to improve the AHI and upper airway muscle tone. This is the first report in the literature to show significant correlations between the increase in tone, as measured by the IOPI, and improvements in the AHI in patients with severe OSAHS.

The main criticisms of previous reports on the use of myofunctional therapy to treat OSAHS patients are that the myofunctional therapy included different kinds of exercises and the authors did not explain the reason for improvement [8]. The clinical trial by Guimaraes et al [10] reported a significant reduction in neck circumference, which may explain the effectiveness of the exercises. Our app provides one unique set of exercises with acoustic, visual, and tactile feedback. We found no significant change in neck circumference, which suggests that increasing muscle tone, as measured by the IOPI, helped to improve the AHI.

Eckert [20] reported the use of myofunctional therapy and UASDs as an elective treatment for patients with a hypotonic phenotype associated with OSAHS. Our results may be similar to those obtained with a stimulator [21], but the app costs less and does not appear to be associated with adverse effects. Another difference after myofunctional therapy is that the IOPI scores increased after use of the app, whereas UASDs have not been found to increase IOPI scores [22]. Active voluntary movement is required to increase the muscle tone, but UASDs do not elicit such movement. However, the app must be evaluated for a longer time, 1 year or more, as performed with UASDs, to make definitive conclusions [21].

We believe that the reason for the success of this app in patients with severe OSAHS is that myofunctional therapy must be based on proprioceptive training because of the sensorimotor deficit in the upper airway muscles in these patients [4]. All clinical studies reported to date are based on isometric and isotonic exercises in patients with moderate or severe sleep apnea, and used videos and diagrams [10,13,19]. Our app is based on proprioceptive training using isometric and isotonic exercises, which led to satisfactory results in these patients with severe OSAHS.

The main disadvantages of myofunctional therapy are the poor adherence to therapy [23] and the absence of objective feedback [10,11]. Adherence to myofunctional therapy was high in our study, possibly because of the ease of contact with a practitioner and the acoustic and visual feedback about the patient's performance given by the app for every exercise [24]. Monthly visits for measurement of IOPI scores can also provide an objective way to evaluate patient progress and to promote adherence.

Instead of using placebo exercises, as employed in other studies [25], we chose to withhold therapy for the control group because

there is no separate app that could be used as a control. This study also describes a new method for delivering upper airway exercise training for which there is no comparable study available to date.

The selection of patients suitable for myofunctional therapy is important [18]. The efficacy of this therapy may be limited in patients with restricted tongue movement, permanently blocked nose, or temporomandibular joint disorder. Therefore, patients with OSAHS should be examined by a therapist with knowledge of the use of myofunctional therapy before beginning such therapy [26]. In our study, 10 patients (16.6%) were rejected because of one of these conditions. This app is suitable only for improving upper airway muscle tone and does not correct other myofunctional therapy disorders [27]. Future studies are needed to identify the hypotonic phenotype(s) suitable for this therapy. As we have suggested previously [28], patients with a low IOPI score are good candidates for myofunctional therapy.

Although further evidence of the efficacy of this app is needed, we consider that this therapy will help to improve adherence to other treatments, as has been suggested previously [28-30]. In addition, conventional treatments such as CPAP or sleep surgery have been restricted during the COVID-19 pandemic [31], and this app may be a reasonable therapeutic option, as we have reported recently [24].

Despite the strengths of this pilot study mentioned above, we note several limitations. First, the number of participants was small. Second, we found a significant loss of participants in the control group (50%), who were instructed not to perform any therapy once they had been diagnosed with severe OSAHS. Despite this loss, our sample size was similar to that included in other clinical studies of this therapy [10,18,19]. The number of patients in the control group was low because we found significant differences in the therapy group early during the study. We therefore decided not to enroll more patients in the control group because of the difficulties experienced by patients with severe OSAHS not given appropriate therapy and the restrictions on sleep studies during the COVID-19 pandemic. Third, as this was a nonsponsored pilot trial, we initially attempted to perform polysomnography studies in all patients, but this test slowed our trial. We decided to use laboratory polygraphy because we found no significant differences in the accuracy of the outcomes AHI and ODI between polysomnography and polygraphy provided that the pre- and postintervention measures were obtained by the same team and equipment. Finally, this pilot trial was registered at clinicaltrials.gov during the performance of the study.

Conclusions

In patients with OSAHS who performed myofunctional therapy exercises using this app, the severity of symptoms decreased and the tone of the upper airway muscles increased after 3 months. This app may represent a promising treatment for OSAHS given its convenience and availability of the mobile phone market.

Conflicts of Interest

CR is the creator of, and has financial interest in, the AirwayGym app. All other authors declare no conflict of interest.

Multimedia Appendix 1
Study protocol.
[DOCX File , 580 KB - mhealth_v8i11e23123_app1.docx]

Multimedia Appendix 2
CONSORT checklist.
[PDF File (Adobe PDF File), 2414 KB - mhealth_v8i11e23123_app2.pdf]

Multimedia Appendix 3
Screenshots from the app.
[PPTX File , 850 KB - mhealth_v8i11e23123_app3.pptx]

Multimedia Appendix 4
Video of exercise 1: chromatic snake.
[MP4 File (MP4 Video), 11722 KB - mhealth_v8i11e23123_app4.mp4]

Multimedia Appendix 5
Video of exercise 2: Snake.
[MP4 File (MP4 Video), 5819 KB - mhealth_v8i11e23123_app5.mp4]

Multimedia Appendix 6
Video of exercise 3: Chameleon up.
[MP4 File (MP4 Video), 8015 KB - mhealth_v8i11e23123_app6.mp4]

Multimedia Appendix 7
Video of exercise 4: Chameleon down.
[MP4 File (MP4 Video), 15004 KB - mhealth_v8i11e23123_app7.mp4]

Multimedia Appendix 8
Video of exercise 5: Tongue left cheek.
[MP4 File (MP4 Video), 7315 KB - mhealth_v8i11e23123_app8.mp4]

Multimedia Appendix 9
Video of exercise 6: Tongue right cheek.
[MP4 File (MP4 Video), 6410 KB - mhealth_v8i11e23123_app9.mp4]

Multimedia Appendix 10
Video of exercise 7: Pressure under the chin.
[MP4 File (MP4 Video), 7033 KB - mhealth_v8i11e23123_app10.mp4]

Multimedia Appendix 11
Video of exercise 8: Left mandibular pressure.
[MP4 File (MP4 Video), 7782 KB - mhealth_v8i11e23123_app11.mp4]

Multimedia Appendix 12
Video of exercise 9: Right mandibular pressure.
[MP4 File (MP4 Video), 7814 KB - mhealth_v8i11e23123_app12.mp4]

References

1. Mendes FA, Marone SAM, Duarte BB, Arenas ACP. Epidemiologic profile of patients with snoring and obstructive sleep apnea in a university hospital. Int Arch Otorhinolaryngol 2014 Apr;18(2):142-145 [FREE Full text] [doi: 10.1055/s-0033-1359309] [Medline: 25992080]
2. Bakker JP, Wang R, Weng J, Aloia MS, Toth C, Morrical MG, et al. Motivational Enhancement for Increasing Adherence to CPAP: A Randomized Controlled Trial. Chest 2016 Aug;150(2):337-345 [FREE Full text] [doi: 10.1016/j.chest.2016.03.019] [Medline: 27018174]
3. Hanning CD. Laryngeal and velopharyngeal sensory impairment in obstructive sleep apnea. Sleep 2005 Oct;28(10):1335. [doi: 10.1093/sleep/28.10.1335] [Medline: 16295220]
4. Guilleminault C, Huang Y, Quo S. Apraxia in children and adults with obstructive sleep apnea syndrome. Sleep 2019 Dec 24;42(12):zsz168. [doi: 10.1093/sleep/zsz168] [Medline: 31581285]
5. Villa MP, Evangelisti M, Martella S, Barreto M, Del Pozzo M. Can myofunctional therapy increase tongue tone and reduce symptoms in children with sleep-disordered breathing? Sleep Breath 2017 Dec;21(4):1025-1032. [doi: 10.1007/s11325-017-1489-2] [Medline: 28315149]
6. Saboisky JP, Butler JE, Gandevia SC, Eckert DJ. Functional role of neural injury in obstructive sleep apnea. Front Neurol 2012;3:95. [doi: 10.3389/fneur.2012.00095] [Medline: 22715333]
7. Aman JE, Elangovan N, Yeh I, Konczak J. The effectiveness of proprioceptive training for improving motor function: a systematic review. Front Hum Neurosci 2014;8:1075. [doi: 10.3389/fnhum.2014.01075] [Medline: 25674059]
8. Camacho M, Certal V, Abdullatif J, Zaghi S, Ruoff CM, Capasso R, et al. Myofunctional Therapy to Treat Obstructive Sleep Apnea: A Systematic Review and Meta-analysis. Sleep 2015 May 01;38(5):669-675 [FREE Full text] [doi: 10.5665/sleep.4652] [Medline: 25348130]
9. Korhan I, Gode S, Midilli R, Basoglu OK. The influence of the lateral pharyngeal wall anatomy on snoring and sleep apnoea. J Pak Med Assoc 2015 Feb;65(2):125-130. [Medline: 25842544]
10. Guimarães KC, Drager LF, Genta PR, Marcondes BF, Lorenzi-Filho G. Effects of oropharyngeal exercises on patients with moderate obstructive sleep apnea syndrome. Am J Respir Crit Care Med 2009 May 15;179(10):962-966. [doi: 10.1164/rccm.200806-981OC] [Medline: 19234106]
11. O'Connor RC, Garcia-Iriarte M, Casado-Morente J. New app "Apnea Bye" increases adherence in myofunctional therapy to treat sleep disordered breathing. In: Otolaryngol Neck Surg (1 Suppl). 2018 Presented at: American Academy of Otolaryngology-Head and Neck Surgery 2018 Annual Meeting & OTO Experience; October 7-10, 2018; Atlanta, GA. [doi: 10.1177/0194599818787193j]
12. Isetta V, Torres M, González K, Ruiz C, Dalmases M, Embid C, et al. A New mHealth application to support treatment of sleep apnoea patients. J Telemed Telecare 2017 Jan;23(1):14-18. [doi: 10.1177/1357633X15621848] [Medline: 26672606]

13. Camacho M, Robertson M, Abdullatif J, Certal V, Kram YA, Ruoff CM, et al. Smartphone apps for snoring. J Laryngol Otol 2015 Oct;129(10):974-979. [doi: 10.1017/S0022215115001978] [Medline: 26333720]
14. 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]
15. Berry RB, Gamaldo CE, Harding SM, Brooks R, Lloyd RM, Vaughn BV, et al. AASM Scoring Manual Version 2.2 Updates: New Chapters for Scoring Infant Sleep Staging and Home Sleep Apnea Testing. J Clin Sleep Med 2015 Nov 15;11(11):1253-1254. [doi: 10.5664/jcsm.5176] [Medline: 26446251]
16. Marchesan IQ. Lingual frenulum protocol. Int J Orofacial Myology 2012 Nov;38:89-103. [Medline: 23367525]
17. Frequently asked questions. AirwayGym. 2018. URL: https://airwaygym.app/en/faq [accessed 2020-10-20]
18. Ieto V, Kayamori F, Montes MI, Hirata RP, Gregório MG, Alencar AM, et al. Effects of Oropharyngeal Exercises on Snoring: A Randomized Trial. Chest 2015 Sep;148(3):683-691. [doi: 10.1378/chest.14-2953] [Medline: 25950418]
19. Torres-Castro R, Vilaró J, Martí JD, Garmendia O, Gimeno-Santos E, Romano-Andrioni B, et al. Effects of a Combined Community Exercise Program in Obstructive Sleep Apnea Syndrome: A Randomized Clinical Trial. J Clin Med 2019 Mar 14;8(3):361 [FREE Full text] [doi: 10.3390/jcm8030361] [Medline: 30875753]
20. Eckert DJ. Phenotypic approaches to obstructive sleep apnoea - New pathways for targeted therapy. Sleep Med Rev 2018 Feb;37:45-59. [doi: 10.1016/j.smrv.2016.12.003] [Medline: 28110857]
21. Strollo PJ, Soose RJ, Maurer JT, de Vries N, Cornelius J, Froymovich O, STAR Trial Group. Upper-airway stimulation for obstructive sleep apnea. N Engl J Med 2014 Jan 09;370(2):139-149. [doi: 10.1056/NEJMoa1308659] [Medline: 24401051]
22. Wirth M, Unterhuber D, von Meyer F, Hofauer B, Ott A, Edenharter G, et al. Hypoglossal nerve stimulation therapy does not alter tongue protrusion strength and fatigability in obstructive sleep apnea. J Clin Sleep Med 2020 Feb 15;16(2):285-292. [doi: 10.5664/jcsm.8184] [Medline: 31992396]
23. Díaz S M, Salazar C A, Bravo G F, Ocampo-Garcés A. Tratamiento del síndrome de apneas e hipopneas obstructivas del sueño con terapia miofuncional orofaríngea: Experiencia en hospital público de Chile. Rev Otorrinolaringol Cir Cabeza Cuello 2019 Dec;79(4):395-403. [doi: 10.4067/s0718-48162019000400395]
24. O'Connor Reina C, Plaza G, Ignacio-Garcia JM, Baptista Jardin P, Garcia-Iriarte MT, Casado-Morente JC, et al. New mHealth application software based on myofunctional therapy applied to sleep-disordered breathing in non-compliant subjects. Sleep Science Practice 2020 Feb 5;4(1):3. [doi: 10.1186/s41606-019-0040-8]
25. Cunali PA, Almeida FR, Santos CD, Valdrichi NY, Nascimento LS, Dal-Fabbro C, et al. Mandibular exercises improve mandibular advancement device therapy for obstructive sleep apnea. Sleep Breath 2011 Dec;15(4):717-727. [doi: 10.1007/s11325-010-0428-2] [Medline: 20967571]
26. Guilleminault C, Huseni S, Lo L. A frequent phenotype for paediatric sleep apnoea: short lingual frenulum. ERJ Open Res 2016 Jul;2(3):00043-2016 [FREE Full text] [doi: 10.1183/23120541.00043-2016] [Medline: 27730205]
27. Mason RM, Franklin H. Orofacial Myofunctional Disorders and Otolaryngologists. Otolaryngology 2014;04(04):e110. [doi: 10.4172/2161-119x.1000e110]
28. Diaféria G, Santos-Silva R, Truksinas E, Haddad FLM, Santos R, Bommarito S, et al. Myofunctional therapy improves adherence to continuous positive airway pressure treatment. Sleep Breath 2017 May;21(2):387-395. [doi: 10.1007/s11325-016-1429-6] [Medline: 27913971]
29. Aardoom JJ, Loheide-Niesmann L, Ossebaard HC, Riper H. Effectiveness of eHealth Interventions in Improving Treatment Adherence for Adults With Obstructive Sleep Apnea: Meta-Analytic Review. J Med Internet Res 2020 Feb 18;22(2):e16972 [FREE Full text] [doi: 10.2196/16972] [Medline: 32130137]
30. Shin JC, Kim J, Grigsby-Toussaint D. Mobile Phone Interventions for Sleep Disorders and Sleep Quality: Systematic Review. JMIR Mhealth Uhealth 2017 Sep 07;5(9):e131 [FREE Full text] [doi: 10.2196/mhealth.7244] [Medline: 28882808]
31. Grote L, McNicholas WT, Hedner J, ESADA collaborators. Sleep apnoea management in Europe during the COVID-19 pandemic: data from the European Sleep Apnoea Database (ESADA). Eur Respir J 2020 Jun;55(6):2001323 [FREE Full text] [doi: 10.1183/13993003.01323-2020] [Medline: 32366489]

Abbreviations

AHI: apnea–hypoxia index
CONSORT: Consolidated Standards of Reporting Trials
CPAP: continuous positive airway pressure
IOPI: Iowa Oral Performance Instrument
mHealth: mobile health
ODI: oxygen desaturation index
OSAHS: obstructive sleep apnea–hypopnea syndrome
UASD: upper airway stimulation device

Edited by G Eysenbach, L Buis; submitted 02.08.20; peer-reviewed by I Bouloukaki; comments to author 24.08.20; revised version received 05.09.20; accepted 22.10.20; published 09.11.20.

Please cite as:
O'Connor-Reina C, Ignacio Garcia JM, Rodriguez Ruiz E, Morillo Dominguez MDC, Ignacio Barrios V, Baptista Jardin P, Casado Morente JC, Garcia Iriarte MT, Plaza G
Myofunctional Therapy App for Severe Apnea–Hypopnea Sleep Obstructive Syndrome: Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(11):e23123
URL: http://mhealth.jmir.org/2020/11/e23123/
doi:10.2196/23123
PMID:33093013

Original Paper

Technology-Enabled Health Care Collaboration in Pediatric Chronic Illness: Pre-Post Interventional Study for Feasibility, Acceptability, and Clinical Impact of an Electronic Health Record–Linked Platform for Patient-Clinician Partnership

Lisa Opipari-Arrigan[1,2,3*], PhD; Dana M H Dykes[4], MD; Shehzad A Saeed[5,6], MD; Sunny Thakkar[1], BA; Lisa Burns[3,7], MD; Barbara A Chini[3,7], MD; Gary L McPhail[3,7], MD; Ian Eslick[8], PhD; Peter A Margolis[1,3], MD PhD; Heather C Kaplan[1,3,9*], MD, MSCE

[1]Anderson Center for Health Systems Excellence, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States
[2]Behavioral Medicine and Clinical Psychology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States
[3]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, United States
[4]GI Care for Kids, Atlanta, GA, United States
[5]Gastroenterology and Nutrition, Dayton Children’s Hospital, Dayton, OH, United States
[6]Gastroenterology and Nutrition, Wright State University, Dayton, OH, United States
[7]Pulmonary Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States
[8]Vital Labs, Inc, San Francisco, CA, United States
[9]Perinatal Institute, Cincinnati Children's Hospital Medical Center, Cincinatti, OH, United States
[*]these authors contributed equally

Corresponding Author:
Lisa Opipari-Arrigan, PhD
Anderson Center for Health Systems Excellence
Cincinnati Children's Hospital Medical Center
3333 Burnet Ave
Cincinnati, OH, 45229
United States
Phone: 1 5136364200
Email: lisa.opipari@cchmc.org

Abstract

Background: Mobile health (mHealth) technology has the potential to support the Chronic Care Model’s vision of closed feedback loops and patient-clinician partnerships.

Objective: This study aims to evaluate the feasibility, acceptability, and short-term impact of an electronic health record–linked mHealth platform *(Orchestra)* supporting patient and clinician collaboration through real-time, bidirectional data sharing.

Methods: We conducted a 6-month prospective, pre-post, proof-of-concept study of *Orchestra* among patients and parents in the Cincinnati Children’s Hospital inflammatory bowel disease (IBD) and cystic fibrosis (CF) clinics. Participants and clinicians used *Orchestra* during and between visits to complete and view patient-reported outcome (PRO) measures and previsit plans. Surveys completed at baseline and at 3- and 6-month follow-up visits plus data from the platform were used to assess outcomes including PRO completion rates, weekly platform use, disease self-efficacy, and impact on care. Analyses included descriptive statistics; pre-post comparisons; Pearson correlations; and, if applicable, effect sizes.

Results: We enrolled 92 participants (CF: n=52 and IBD: n=40), and 73% (67/92) completed the study. Average PRO completion was 61%, and average weekly platform use was 80%. Participants reported improvement in self-efficacy from baseline to 6 months (7.90 to 8.44; *P*=.006). At 6 months, most participants reported that the platform was useful (36/40, 90%) and had a positive impact on their care, including improved visit quality (33/40, 83%), visit collaboration (35/40, 88%), and visit preparation (31/40, 78%). PRO completion was positively associated with multiple indicators of care impact at 3 and 6 months.

Conclusions: Use of an mHealth tool to support closed feedback loops through real-time data sharing and patient-clinician collaboration is feasible and shows indications of acceptability and promise as a strategy for improving pediatric chronic illness management.

(JMIR Mhealth Uhealth 2020;8(11):e11968) doi:10.2196/11968

KEYWORDS

health services research; mHealth; eHealth; patient engagement; chronic illness; mobile phone

Introduction

Background

Optimal management of pediatric chronic illness requires a different type of health care system [1,2]. Current models of care organized around treating acute, episodic conditions no longer meet the needs of the growing number of children with chronic health problems [3,4]. Although, organizationally, care is focused around the clinic visit, more than 99% of patients' lives are spent outside of this context. The Chronic Care Model positions *informed, activated patients* and *prepared, proactive clinicians* as foundational to improved health outcomes. However, it is difficult for most clinicians to be informed about what happens outside of the clinic. In turn, most patients have neither easy, understandable access to their own clinical data nor systematic ways to reliably measure and communicate changes in daily functioning. This is a missed opportunity for optimizing health care. Failing to connect what happens day-to-day with clinical decision-making leads to reactive instead of proactive care based on hazy memories, missing information, and subjective impressions of *how things have been going.*

Mobile health (mHealth) technology offers the opportunity to address these gaps by expanding care beyond the boundaries of time-limited, infrequent clinical encounters to support data sharing, communication, and collaborative decision-making between patients and clinicians [5-9]. This potential has encouraged considerable attention from the health care and technology industries; however, we have yet to see mHealth have a transformative effect on chronic care delivery systems or outcomes. Although the number of digital health apps, portals, and dashboards is rapidly increasing, evidence of their impact has been mixed and modest [10]. Some studies have demonstrated that mHealth supports improved engagement and adherence. For example, text messaging interventions have boosted attendance at pediatric HIV appointments [11] and have increased adherence in pediatric asthma [12], type 1 diabetes [13], and cardiac care [14]. Electronic health portals that track patients' symptoms and provide decision support have decreased clinic visits and school absences in pediatric ulcerative colitis [15] and decreased flares and missed parental work in pediatric asthma [16]. Yet, there are countless other examples of mHealth technologies that fail to improve self-management behaviors or clinical outcomes [9,10,16-21]. In fact, a comprehensive review of mHealth technology for chronic disease management found that just over half impacted adherence behavior and only 39% impacted clinical outcomes [10]. In addition, studies show that about half of the individuals that download health apps stop using them, most soon after initial use [22,23].

Both technological design elements and clinical care integration likely contribute to an mHealth tool's success or failure [24]. Tools that only enable tracking and display of user data may build awareness but are likely insufficient to engage individuals in health behavior change [24-26] and surmount the burden of data entry [22]. Rather, mHealth tools that support the Chronic Care Model's vision of closed feedback loops and patient-clinician partnership, through making data relevant, understandable, and actionable and by facilitating collaboration, are more likely to be adopted and retained [26,27]. Beyond the technology itself, purposeful integration with clinical care, so that both patients and clinicians understand the *why* and *how* of incorporating mHealth tools, is paramount to achieving impact [10,26,28].

Objectives

The purpose of this study is to test the feasibility, acceptability, and short-term impact of using an mHealth tool to facilitate the Chronic Care Model's vision of closed feedback loops and patient-clinician partnerships. As there was no available commercial platform that reliably enabled the functionality needed to support the Chronic Care Model's approach to continuous care, we developed an mHealth tool with patient and clinician end users to support pediatric chronic care during and between clinical encounters. The *Orchestra* platform [29] was designed to enable proactive and collaborative chronic care through (a) real-time sharing and visualization of clinical and patient-generated (ie, patient-reported outcome, PRO) data, (b) automated symptom surveillance with actionable alerts to signal changes in patient status based on patient-generated data, and (c) collaborative previsit planning. As a first step, this prospective, preliminary pre-post study tested *Orchestra* in 2 pediatric chronic diseases (cystic fibrosis [CF] and inflammatory bowel disease [IBD]). We hypothesized that parents, patients, and clinicians would find *Orchestra* feasible and acceptable and that we would see signals of clinically meaningful impact.

Methods

Setting and Patient Population

This study was conducted within the IBD and CF clinics at Cincinnati Children's Hospital (CCH; Ohio, United States) and was approved by the CCH Institutional Review Board (#2014-0975). We conducted the work at CCH because of our existing relationships with clinics that had an interest and ability to partner with us in this proof-of-concept study. We used a convenience sample of patients and parents from the CCH CF and IBD clinics. The CF and IBD clinics were chosen to ensure that our study included conditions that varied in disease course, management, and time demands associated with daily care.

Eligible participants were patients with IBD or CF aged between 14 and 21 years and/or a parent or legal guardian of patients aged between 0 and 21 years. If a patient was less than 14 years old, only their parent was enrolled in the study and used the *Orchestra* app. If a patient was older than 14 years, the family decided if only the patient, only the parent, or both would enroll and use the *Orchestra* app. Participants were English speaking or reading and had a smartphone (iOS or Android) with a mobile data plan and/or internet connectivity.

Study Design and Protocol

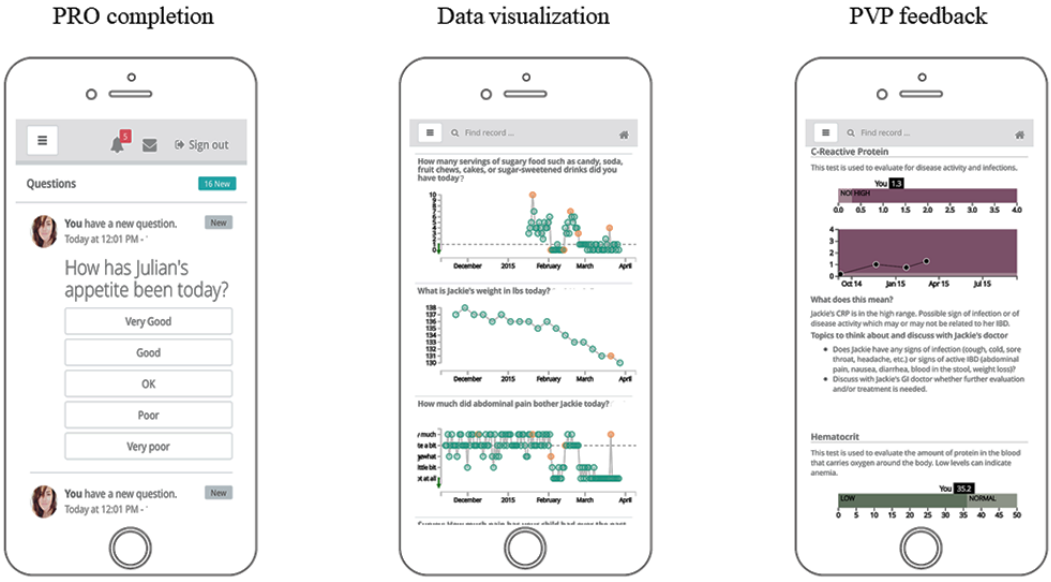
This was a preliminary, prospective pre-post, proof-of-concept study. Recruitment began in May 2015 and continued for 1 year. Follow-up data collection was completed in fall 2016. Eligible participants were identified via clinic rosters and were contacted by research staff, who were not part of the clinical team, before or at a clinic visit. Interested participants watched a 3-min video [30] describing *Orchestra* while waiting for their visit and then discussed with their clinicians how they might use *Orchestra* as part of their care. Following the visit, participants met with staff to sign the informed consent form, download the *Orchestra* mobile app, and customize settings. Participants were enrolled for 6 months.

Orchestra Technology Platform

The *Orchestra* technology platform [29] was developed with clinicians, patients, and parents via a user-centered, agile design process. The platform included a patient- or parent-facing mobile app and a linked clinician-facing web-based dashboard. Patient or parent functionality included the following (Figure 1):

- Library of PRO measures [31,32] that could be selected for symptom and general health tracking based on patient goals and customized for daily or weekly completion. There were 35 predetermined measures available for IBD (eg, abdominal pain frequency, stool consistency, and fatigue) and 28 available for CF (eg, energy level, cough severity, and appetite). Participants could also create custom symptoms with their clinician if needed (Multimedia Appendix 1).
- Real-time visualization of PRO data.
- PRO data point annotation and journal entry features for recording observations.
- Previsit health report showing disease-specific laboratory data (ie, C-reactive protein and hematocrit) and health metrics (ie, weight and pulmonary function tests) with simple visualizations and text descriptions of results and trends as well as personalized suggestions for topics to discuss with clinicians.
- Previsit plan (PVP) completion to inform clinical teams about patient symptoms, goals, and questions before a clinic visit.

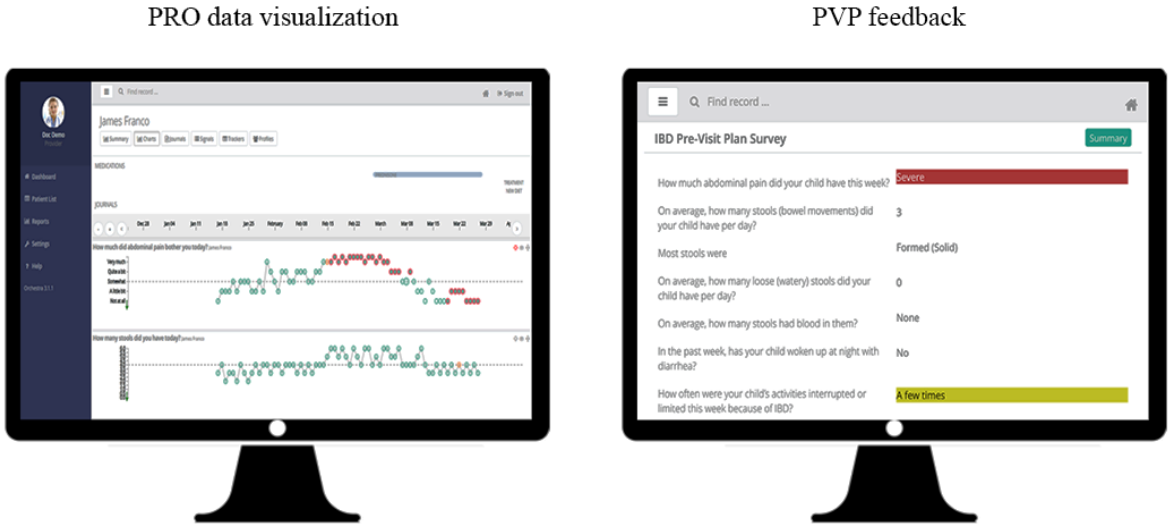
Figure 1. "Orchestra" participant mobile app example screens. PRO: patient-reported outcome; PVP: previsit plan.



The clinician-facing web-based dashboard was accessed via Epic (Epic Systems Corporation), the electronic medical record, using a customized configuration that enabled one-click, single sign-on authentication. Clinical users included physicians and nurses in IBD and physicians, nurses, dietitians, a social worker, a respiratory therapist, and a psychologist in CF. The clinician dashboard enabled the following (Figure 2):

- Real-time visualization of PRO data and review of PVPs.
- Addition of notes to a patient's account visible by the patient or parent and other clinical staff.
- Notification if a patient's symptoms met prespecified criteria based on PRO data. Alert triggers were customizable and based on meeting a threshold value and/or meeting standard statistical process control criteria of 8 data points above or below the median after a 20-data point baseline was established [33]. If alerts were triggered, clinicians were notified via the *Orchestra* dashboard and email.

Figure 2. "Orchestra" clinician dashboard example screens. PRO: patient-reported outcome; PVP: previsit plan.



Orchestra Intervention

Use of the *Orchestra* technology platform was accompanied by a delivery system intervention to integrate the technology into clinical workflows and patients' lives, personalize goals for use, and establish accountability for system use between participants and clinicians. All clinicians who cared for patients, including physicians, nurses, dietitians, a respiratory therapist, a social worker, and a psychologist, received training on the use of *Orchestra* [29]. Participants and clinicians used shared decision-making during the baseline clinic visit to discuss their preferences and goals for *Orchestra* use, select which PROs to complete and at what frequency, and decide who (patient and/or parent) would use the tool. Clinicians discussed health concerns and goals with patients (eg, reducing symptoms, reducing medications, and preventing exacerbations) and identified whether using *Orchestra* would help achieve these goals. The clinician and family then reviewed the PRO and symptom measures available in *Orchestra* and determined which measures were most relevant to addressing the identified goals. They also established a *social contract* to set expectations regarding between-visit data completion (by participant), review (by clinician), and communication (participant-clinician). Clinicians chose whether to set data alerts and customized trigger points based on the PROs selected and patient needs.

To reinforce the use of *Orchestra* by demonstrating that data collected between visits were reviewed, a designated clinical team member was asked to review the data entered into the *Orchestra* platform each week and within 1 day of receiving a data alert. In addition, clinicians were asked to review PVP and PRO data before all clinic visits. At visits, clinicians were encouraged, at least, to verbally acknowledge reviewing the data and ideally to use the data for discussion and decision making.

Data Collection

Participants completed study assessments at baseline and 3- and 6-month follow-up visits during regularly scheduled clinic appointments. As this tool was designed to be used as part of clinical care, if a patient did not have a clinic appointment within ±1 month of a follow-up time point, the assessment was not completed. Participants were still included in the study if they did not have an appropriately timed follow-up visit, but that specific assessment was considered missing.

Health status indicators, including the Physician Global Assessment (PGA) Score (IBD) and forced expiratory volume in 1 second (FEV1) percent predicted (CF), were collected from the electronic health record to understand the baseline health of participants. The PGA score is an indicator of disease severity that is rated at each clinic visit by the physician and is based on symptoms, clinical exam, and laboratory measures [34]. PGA is used to classify disease as in remission, mild, moderate, or severe. FEV1 is calculated as part of pulmonary function testing and is an established marker of disease progression in CF [35]. The lower the FEV1 percent predicted, the more severe the lung disease.

Clinicians and research staff reported on baseline technology discussion and set-up time via a paper and pencil measure completed immediately following the patient visit. In addition, the patient's physician and, if applicable, a second clinical team member involved in using Orchestra completed surveys after each visit regarding their assessment of the patient and parent's engagement in the visit and their experience using Orchestra. However, due to an inability to obtain responses in a timely manner from clinicians, the data were not analyzed.

Outcomes

The *feasibility* of implementing and supporting *Orchestra* in care was assessed by measuring the percentage of eligible patients with access to a smartphone with data plan or internet connectivity, minutes spent discussing *Orchestra* (self-reported by the clinician at baseline), minutes spent enrolling and setting up the *Orchestra* app (self-reported by research staff), and participants' report of whether *Orchestra* was used during follow-up visits. One facet of *acceptability* was assessed by measuring usage of the app including PRO completion (ie, percentage of PROs a participant answered), weekly engagement with the mobile app (ie, percentage of weeks a participant input data at least once), and PVP completion (percentage of PVPs

that were submitted in the app). We examined the short-term *impact* of *Orchestra* on disease self-efficacy, visit engagement, and care impact, including perceived care quality, patient-clinician collaboration, visit preparation, disease insight, treatment plan quality, and perceived usefulness. As we did not expect health outcomes to change over the 6-month study period, we focused on measures of short-term meaningful clinical impact that reflected participants' perceptions of their ability to manage their illness (efficacy) as well as their perceptions of their clinic visit and interaction with their clinicians. Disease self-efficacy was measured using the validated Self-Efficacy for Managing Chronic Disease 6-Item Questionnaire [36,37]. Each item uses a 1 (*not at all confident*) to 10 (*totally confident*) scale, and a mean of all items is calculated. Visit engagement was measured using an 8-item survey examining visit preparation, knowledge, and involvement given at baseline and at the 3- and 6-month follow-up (Multimedia Appendix 2). As one item assessed the use of tracked data at clinic visits, it was assessed only at 3 and 6 months and was not included in the analysis. At 3 and 6 months only, participants completed a 20-item care impact survey that measured the impact of *Orchestra* on care quality (4 items), patient-clinician collaboration (2 items), visit preparation (6 items), disease insight (3 items), treatment plan quality (3 items), and perceived usefulness (2 items; Multimedia Appendix 2). Both the visit engagement and care impact surveys were developed theoretically by a team of experts for this study, used a 6-point scale ranging from 1 (*strongly disagree)* to 6 (*strongly agree)*, and were tested in a convenience sample for understandability. The 3- and 6-month surveys also included a single question measuring overall satisfaction: "How likely are you to recommend Orchestra to other parents/patients like you?", which respondents answered on a 10-point scale from 10 (*very likely*) to 1 (*not at all likely*). Impact on the treatment plan was measured by chart review examining the percentage of data alerts that led to documented clinical action including phone calls, emails, or messages to check status; expedited clinic appointments or laboratory tests; referrals; or treatment changes.

Statistical Analyses

Descriptive statistics, including means and ranges for continuous variables and percentages for categorical variables, were used to describe the study population, feasibility, acceptability, satisfaction, and perceived impact. A mean score of 4 or higher on the visit engagement survey and care impact survey subscales was defined as an indicator of engagement in the visit and improvement in care experience, respectively. *t* tests were used to assess the impact of condition (CF or IBD), person using *Orchestra* (parent or patient), and baseline health status in CF (FEV1: ≥70% predicted or <70% predicted) on PRO completion. Cohen *d* measured the effect size. Univariate analysis of variance tests was used to evaluate differences in acceptability measures across study clinicians and across IBD baseline health status (PGA: continuous remission, remission, mild, and moderate). Paired *t* tests examined the impact on disease self-efficacy and visit engagement from baseline to 6-month follow-up. Pearson correlation coefficients examined the relationships between PRO completion and perceived impact on care. Analyses were conducted using the Statistical Package for the Social Sciences v23 (SPSS, IBM Corp).

We aimed to enroll 100 participants. Our target was based on knowledge of the eligible population in each clinic, frequency of clinic visits, available research resources, and a desire to complete the study while the technology remained current. As efficacy was not the primary endpoint, we did not perform an a priori power calculation.

Results

Overview

Of the 127 families approached, 88 (69.2%) agreed to participate. A total of 92 participants (CF: n=52 and IBD: n=40) from 88 families enrolled in the study (Figure 3). In 38 families (38/88, 43%), an adolescent patient participated; in 46 families (46/88, 52%), a parent participated; and in 4 families (4/88; 5%), both participated. At 3 months, 71 of the original 92 participants (77%) remained in the study; at 6 months, 67 of the original 92 participants (73%) remained (Figure 3). The 6-month retention rate was higher among IBD clinic participants (32/40, 80%) than among CF clinic participants (35/52, 67%). Eight different attending physicians had participants enrolled on the platform, 2 from IBD and 6 from CF. This included all attending physicians in the CF clinic and the physicians who most frequently saw patients with IBD in the gastroenterology clinic. Table 1 describes the demographic and health status characteristics of the participants. Characteristics of the participants mirrored the general demographics of the clinic populations. Participants selected an average of 4.96 PROs (range: 1-13) to track in the app.

Figure 3. Participant flow diagram. Of the 127 eligible families, 88 were enrolled (92 participants). At 3 months, 77% (71/92) of participants remained in the study and 66% (47/71) completed the 3-month follow-up within the specified study time window. At 6 months, 74% (67/92) of the participants remained in the study and 63% (42/67) completed the 6-month follow-up within the specified study time window.

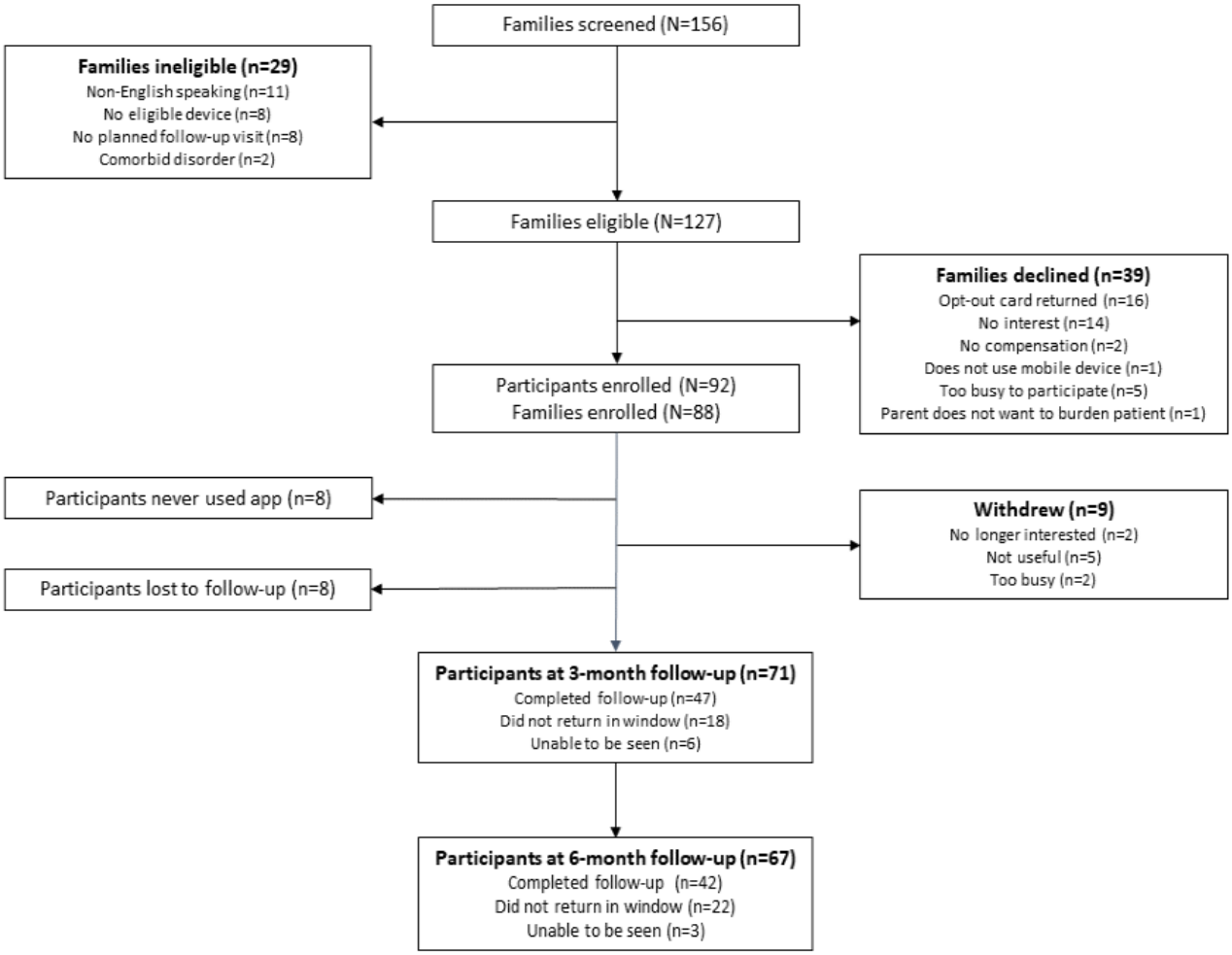


Table 1. Baseline demographic and health status characteristics of the patients for each enrolled family.

Patient characteristics	Cystic fibrosis (n=51)	Inflammatory bowel disease (n=37)	Full sample (N=88)
Patient age (years), mean (SD)	10.55 (6.01)	15.70 (4.02)	12.72 (5.82)
0-2, n (%)	7 (14)	1 (3)	8 (9)
3-12, n (%)	24 (47)	9 (24)	33 (38)
12-21, n (%)	20 (39)	27 (73)	47 (53)
Patient gender, n (%)			
Female	24 (47)	22 (60)	46 (52)
Male	27 (53)	15 (40)	42 (48)
Patient race, n (%)			
White	49 (96)	35 (95)	84 (96)
Black or African American	0 (0)	2 (5)	2 (2)
Other	2 (4)	0 (0)	2 (2)
Disease indicators, mean (SD)			
Mean forced expiratory volume in 1 second percent predicted[a,b] (cystic fibrosis only, n=38)	96.89 (19.69)	—[c]	—[c]
Disease type (inflammatory bowel disease only)	—[d]	• Crohn: 29 (78%) • Ulcerative Colitis: 8 (22%)	—[d]
Physicians global assessment score (inflammatory bowel disease only)	—[d]	• Continuous Remission: 10 (27%) • Remission: 14 (38%) • Mild: 11 (30%) • Moderate: 2 (5%)	—[d]
Family income (US $), n (%)			
<25,000	4 (8)	3 (8)	7 (8)
25,000 to 49,999	12 (24)	1 (2)	13 (15)
50,000 to 74,999	8 (15)	5 (14)	13 (15)
75,000 to 99,999	9 (18)	7 (19)	16 (18)
≥100,000	15 (29)	14 (38)	29 (33)
Unknown	3 (6)	7 (19)	10 (11)

[a]Scores ranged between 48 and 127.

[b]Not all participants were old enough to reliably complete pulmonary function testing.

[c]Measure not applicable for patients with inflammatory bowel disease.

[d]Measure not applicable for patients with cystic fibrosis.

Feasibility

Access to technology was not a significant barrier to participation. Of the families approached, 119 of 127 (94%) had regular access to an eligible device and internet connectivity. Most participant-clinician conversations about *Orchestra* (40/83, 48%) required 5-10 min, with 34% (28/83) taking <5 min and 18% (15/83) requiring >10 min. *Orchestra* enrollment and app installation took 14 min on average (SD 7; 95% CI 13-16), ranging between 2 and 43 min (data available for n=84 families). Longer set-up times were related to technical issues or need for more thorough instruction. It was feasible to incorporate *Orchestra* into follow-up clinic visits—participants reported discussing *Orchestra* in 83% (34/41) of visits at 3 months and in 71% (29/41) of visits at 6 months.

Acceptability

Among those who used the PRO feature (n=80), 19,954 PROs were completed. The average PRO completion rate by participants was 61% (range: 6%-100%; 95% CI 55-68). There was no difference in mean PRO completion rates between CF (n=47, 58%; 95% CI 50-67) and IBD (n=33, 66%; 95% CI 58-75; *P*=.19). Parents (n=41) had significantly higher rates of PRO completion (67%; 95% CI 59-76) than patients (n=39; 55%; 95% CI 46-64; *P*=04; Cohen *d*=0.46). Participant PRO completion rates varied significantly across different physicians (n=8), ranging between 38% and 77% (*P*=.02; partial eta squared=0.20). PRO completion was not related to baseline health status in CF (FEV1 percent predicted) or IBD (PGA score), data not shown.

Average weekly engagement, a measure of app *stickiness* that indicates the participants' willingness to return to the app, was 80% (n=80; range: 21%-100%; 95% CI 75-85). No significant differences in weekly engagement were observed between CF (n=47; 80%; 95% CI 73-87) and IBD (n=33; 80%; 95% CI 72-88; *P*=.99) or based on participants' physicians (n=8; *P*=.06). Parents (n=41) had a higher rate of weekly engagement (87%; 95% CI 81-92) compared with patients (n=39; 73%; 95% CI 65-81; *P*=.005; Cohen *d*=0.64). Patients with IBD in continuous remission for ≥1 year based on PGA (n=6) showed lower weekly engagement (55%; 95% CI 28-83) than patients in current remission for <1 year (n=13; 84%; 95% CI 71-96) or those rated as having mild (n=12; 86%; 95% CI 75-98]) or moderate (n=2; 93%; 95% CI 2-184) disease activity (*P*=.02; partial eta squared=0.28). No differences were found in the average weekly engagement based on FEV1 status in CF (n=35).

Participants received three PVPs on average (range: 0-6) and completed 49.4% (94/190). Most families who received a PVP (56/76, 74%) completed the information at least once.

Short-Term Impact

Participants reported significant improvement in disease self-efficacy from baseline (7.90, SD 1.63; 95% CI 7.70-8.34) to the 6-month follow-up (n=40; 8.44, SD 1.34; 95% CI 8.01-8.87; *P*=.006; Cohen *d*=0.36). The majority of participants reported feeling engaged in their clinic visit at baseline (87/92, 95%) and 6-month follow-up (38/40, 95%), and there was no change in self-reported visit engagement scores from baseline (5.13, SD 0.74; 95% CI 4.84-5.22) to 6-month follow-up (n=40; 5.29, SD 0.66; 95% CI 5.06-5.50; *P*=.12). Most participants reported that incorporating *Orchestra* into care was useful and had a positive impact on care quality, visit collaboration, participant visit preparation, participant disease insight, and treatment plan quality. Table 2 shows the percentage of participants who responded that *Orchestra* improved their care experience. Participants who used *Orchestra* more were more likely to report greater perceived impact on care. As shown in Table 3, the higher a participant's PRO completion, the more useful they found *Orchestra* and the more impact it had on reports of care quality, visit collaboration, participant visit preparation, and treatment plan quality at 3- and 6-month follow-up visits.

Table 2. Percentage of Orchestra users reporting that the tool improved their care experience.

Impact on care experience	3-Month follow-up (n=44)[a], n (%)	6-Month follow-up (n=40)[b], n (%)
Improved care quality	40 (91)	33 (83)
Improved collaboration	37 (84)	35 (88)
Improved participant preparation for visit	35 (80)	31 (78)
Improved participant disease insight	33 (75)	29 (73)
Improved treatment plan	38 (86)	32 (80)
Perceived useful for visit	40 (91)	36 (90)

[a]44 out of 47 participants completed the care impact survey at the 3-month follow-up.

[b]40 out of 42 participants completed the care impact survey at the 6-month follow-up.

Table 3. Pearson correlation coefficients between patient-reported outcome completion rate and impact of Orchestra on care experience.

Impact on care experience	3-Month follow-up (n=43)[a]	*P* value	6-Month follow-up (n=39)[b]	*P* value
Improved care quality	0.52	<.001	0.33	.04
Improved collaboration	0.39	.009	0.38	.02
Improved participant preparation for visit	0.47	.002	0.34	.04
Improved participant disease insight	0.33	.03	0.07	.67
Improved treatment plan	0.51	<.001	0.36	.03
Perceived useful for visit	0.51	<.001	0.22	.18

[a]44 out of 47 participants completed the care impact survey at the 3-month follow-up, and 43 out 47 participants received regular patient-reported outcomes (n=1, used tool for previsit plan only).

[b]40 out of 42 participants completed the care impact survey at the 6-month follow-up, and 39 out of 42 participants received regular patient-reported outcomes (n=1, used tool for previsit plan only).

Participants reported moderately high satisfaction (ie, likelihood of recommending to others) with using *Orchestra* at the 3-month (mean 8.16, SD 1.76; n=45) and 6-month follow-ups (mean 8.55, SD 2.05; n=40).

Clinicians set optional alerts for 49 of 92 (53%) participants. A total of 192 alerts were triggered among 35 of the 49 (71%) patients. On the basis of the clinician report and medical record review, 39% (74/192) of alerts led to documented, actionable clinician behavior. For example, in one case, the CF team was signaled about a change in mucous production and cough based

on *Orchestra* alerts. This prompted a phone call to the family who had noted the symptoms but had not acted on them. The clinical team was able to initiate early treatment (antibiotics and increased airway clearance) and potentially avoid an inpatient admission for pulmonary exacerbation. More than half (65/111, 59%) of alerts in IBD led to action as compared with only 11% (9/81) in CF.

Discussion

Principal Findings

We tested a technology platform and care delivery system intervention designed to support the Chronic Care Model's vision of informed patients and prepared clinicians working in partnership with closed feedback loops. This proof-of-concept study suggests that the mHealth platform is feasible, acceptable, and shows promise for clinically meaningful impact. The average PRO completion rate by participants across the 6 months of the study was 61%, and the average weekly platform use was 80%, suggesting reasonable acceptability. Participants reported significant improvement in self-efficacy from baseline to 6 months. At 6 months, most participants reported that the platform was useful and had a positive impact on their care, including improved visit quality, visit collaboration, and visit preparation. PRO completion rate was moderately associated with participant-reported visit preparation, care quality, treatment plan quality, collaboration, disease insight, and usefulness at 3 months. All relationships, except disease insight and usefulness, remained significant, although weaker, at 6 months.

The addition of shared mHealth technology into the process of managing chronic disease enabled frequent, timely, and problem-focused information and communication. We posit that these types of real-time, bidirectional shared data potentiate a fundamental shift in the interaction between clinician and patient—a shift from clinician as the expert delivering care to a relatively passive patient, to clinician and patient contributing to shared work in which the expertise of both is needed. Our findings are consistent with the small number of mHealth studies specifically designed to facilitate patient-clinician collaboration that also show early evidence of effectiveness [16,38] and, as hypothesized, demonstrate that collaboration is associated with sustained use [26]. For example, a pilot study compared pediatric patients with asthma receiving standard care versus access to a portal that included symptom tracking, sharing of concerns and goals, and symptom threshold–based prompts [16]. In this study, more than half of the parents completed >80% of monthly surveys, and they reported improved communication, fewer asthma flares, and less missed work compared with standard care. *Orchestra* achieved similar rates of engagement (80% average weekly engagement) and sustained use (almost 60% of participants engaged for ≥24 weeks). Our work extends these results by demonstrating high rates of use in two other chronic pediatric diseases. We were also able to show that incorporating processes to support patient-clinician collaboration and closed feedback loops, including shared decision-making to establish goals for technology use and reinforcement of PRO completion through regular review and use of the data, is feasible to implement in a busy clinical setting. The benefits of focusing on mHealth users' goals and their commitment to these goals were also underscored in a longitudinal study of mHealth app usage. This study showed that the key to successful use of mHealth technology revolved around persisting at goals while using the right system that fits users' needs [23].

A key strength of our approach was that we focused not only on user-centered technology but also on the processes to support integration into workflow, closed feedback loops, and patient-clinician collaboration. This attention to helping our end users (ie, patients and clinicians) understand the value of between-visit data and feel comfortable with using it was critical to the tool's acceptability and sustained use. Although mHealth apps alone often function as sophisticated data collection and delivery mechanisms, when well-integrated into care, they may support patients and clinicians in moving beyond traditional role expectations to learning to work together in new ways that transform their interactions, health care, and outcomes [39,40].

Given that participants were delivered PROs to complete once per day, the greater than 60% PRO completion rate seen in this study means that participants recorded health data on average about 4 days per week between clinic visits, resulting in a significant amount (almost 20,000 completed PROs) of previously untapped information that was available to inform care. Although technology enables easy data collection, more data are not always better, and it can easily overwhelm clinical teams [9]. Giving clinicians the ability to set data alerts was done as a way to help manage the volume of data in busy clinical workflows by highlighting potentially actionable information. Incorporating statistical process control methods to assist with identifying meaningful changes from the patient's unique baseline was a novel approach. Although alerts were only set for about half of the patients, almost 40% of the alerts that fired led to documented action between clinic visits. Interestingly, more alerts were acted upon in IBD versus CF. Several factors may explain this finding, including more experience among our IBD physicians in selecting and using daily PROs; the complex, multisystem nature of CF; and the sensitivity of the alert threshold chosen by the physicians. Additional work is needed to understand how to balance sensitivity to detect meaningful changes in patient status with the volume of alerts generated. We speculate that the right balance will differ by condition and by patient.

Our data suggested that participants who used *Orchestra* more reported greater perceived impact on care. We found that the higher a participant's PRO completion rate, the more they reported that *Orchestra* had a positive impact on care quality, visit collaboration, participant visit preparation, and treatment plan quality. We hypothesize that PROs, PVP data, real-time visualizations, and alerts better enabled patients and clinicians to work together between and during visits, leading to improvements in short-term outcomes. Specifically, we posit that by responding to daily or weekly PROs, participants had a systematic way to learn, remember, and communicate with clinicians between visits. In turn, the data potentiated clinicians being more informed during clinical contacts and better able to proactively address problems between visits. Another reason participants who completed more PROs may have reported

greater benefits relates to the potential for closed feedback loops that reinforce platform use. Use of shared decision-making at the onset may have increased the likelihood that the data collected had relevance to both the clinician's and participant's goals, creating greater potential for relevant, actionable data. In turn, the potential for relevant, actionable data could have increased the probability that participants entered it and that clinicians reviewed it and talked about it at visits (reported in 83% of visits at 3 months and in 71% of visits at 6 months). As clinicians used tracked data, we believe that participants' efforts were reinforced, leading to sustained data collection and more actionable results. A stronger relationship between tool use and care impact was found at 3 months as compared with 6 months. We hypothesize that this may be related to the length of the intervening time between the goal for tool use being discussed and PROs selected and the subsequent clinic visit. Patients and clinicians determined how and why the tool would be used at baseline only. It may be that 6 months later, the data are less valuable due to essential learning already occurring (perhaps at a prior visit) and/or areas of concern changing. More frequent recalibration between patients and clinicians (ie, every visit) may be needed to sustain the relevance of mHealth tool use in chronic care.

Our findings reveal several factors that may contribute to improved engagement with mHealth technology. Involvement of the clinical team in supporting participants' platform use is critical [26]. Differences in PRO completion rates across physicians and a tendency for participants' weekly engagement to vary by physician is consistent with observations that clinicians varied in skill at helping patients identify goals for tool use and integrating *Orchestra* data into follow-up visits. Understanding how to help clinicians maximize patients' engagement and use of mHealth technology will be important for future interventions. Not surprisingly, parents showed higher levels of overall PRO completion and weekly app engagement as compared with teens. However, teens completed over half of the PROs they were sent. With realistic expectations for app use, this type of mHealth technology could support adolescent care transition. The finding that health status predicted weekly engagement in IBD but not CF suggests that disease course may lead to different patterns of app use. In IBD, participants in continuous remission used the app less than those who were not. Frequent PRO completion during periods of sustained wellness is likely a low-value behavior. In CF, a condition with periods of exacerbation but no remission, app use was not related to health status, suggesting that PRO completion behavior may be influenced by the interplay of disease course and current functioning. Although these two conditions differ significantly in required time for daily self-management, with CF typically being more intensive, average app use across the diseases did not differ, suggesting feasibility even in the context of high disease self-management demands.

Limitations

This study had some limitations. Our evaluation of app use and impact is based on participants' perspectives only. We attempted to collect clinician data on tool use and impact following clinical encounters but were unsuccessful in getting clinicians to complete surveys in a timely manner. Rather than asking physicians to complete the surveys immediately after the visit, we allowed them to answer them later by email. In future work, we will gather the data as soon as the physician leaves the clinic room or consider obtaining clinicians' perspectives via qualitative interviews. As this study focused on proof of concept, we did not include a contemporaneous control group, so it is unclear how those receiving standard care felt about their visits and ability to communicate with their providers in between visits or whether self-efficacy would have gradually improved over time even without the use of *Orchestra*. It is not possible to rule out that either natural history led to the clinical impact observed or that selection bias led to those who were most interested in technology and collaboration choosing to participate, stay in the study, and show benefit. By using limited exclusion criteria, a pragmatic approach that utilized existing clinic visits (rather than research study visits), and studying more than one chronic disease, we attempted to lessen the impact on these potential threats to validity. Although our participants spanned the economic spectrum, one-third of the families included in this study had incomes above US $100,000. Although convenience sampling made sense in this early work, in future studies, using sampling techniques to ensure appropriate representation across the economic spectrum is critical. It will also be important to ensure that this type of technology and intervention does not increase health disparities. Some measures were developed for this study because of the lack of validated instruments to evaluate visit engagement and the impact of the mHealth tool on care. These measures had face validity and were tested for understandability, but we did not conduct psychometric testing to examine factor structure, reliability, or validity. As this was designed as a pragmatic study, we only attempted to collect follow-up data from participants if they had a clinical visit during the follow-up window. As a result, we do not have the follow-up questionnaire data on the full sample at either time point. For example, at 6 months, approximately two-third of the participants (42 of 67) who remained in the study had a clinic visit. There could have been systematic differences between participants who returned to the clinic more and less often. However, given that the intent of the study was as a proof of concept, identifying a tool that may have had a meaningful impact for even a fraction of the patients who used it is beneficial learning.

Lessons Learned

The lessons learned in this process are important for future work on mHealth tools for chronic illness. First, the value of working iteratively and in concert with patient and clinician end users cannot be overstated. Through collaborative development, they showed us their pain points and thus the necessary functionality to sustain engagement. Second, the one-click, single sign-on with the electronic medical record was critical for our clinicians to sustain use. Although for a short time they were excited by the novelty of a new platform, its use was not sustainable in their workflow if it was not reachable from the electronic medical record. Third, the automated data signals were critical in facilitating clinician comfort with between-visit data collection and thus having multiple patients on the platform. Fourth, both clinicians and patients needed time to learn the value of between-visit data. Structuring workflow to encourage

using the data facilitated the learning and was key. Finally, reliable and low-friction technology is necessary, but not sufficient, to support engagement over time. Rather, understanding patients' and clinicians' needs and goals and then using technology to support achieving them is essential for impacting health care and outcomes.

Conclusions

We have shown that the use of an mHealth technology designed to facilitate the Chronic Care Model's vision of closed feedback loops and patient-clinician partnerships is feasible, acceptable, and shows promise as a strategy for improving pediatric chronic illness management. These encouraging early results reinforce the potential of mHealth technology to support collaboration and real-time data sharing. Although the start-up company that developed *Orchestra* is no longer supporting the platform, this proof-of-concept study identifies key components necessary in any mHealth platform designed to support continuous data and patient-clinician collaboration, important processes necessary to include in clinician and patient workflows to facilitate technology use, and lessons for future studies.

Acknowledgments

The authors acknowledge the contributions of all patient and family participants as well as the parent partner, nurses, respiratory therapists, dietitians, social workers, psychologists, and physicians in the CCH IBD and CF clinics. The authors would also like to thank Pamela Schoettker, MS, for her editorial assistance in preparing this manuscript. This study was funded in part by the Robert Wood Johnson Foundation (72313) Enabling Uptake of a Registry-Supported Learning System in the United States. The funder had no role in the design, methods, participant recruitment, data collection, analyses, or preparation of this manuscript.

Conflicts of Interest

Orchestra was developed via a joint effort between Vital Labs, Inc and CCH. IE, HK, PM, and LO were coinventors of Orchestra and are entitled to proceeds from the successful commercialization of the technology; however, Vital Labs, Inc is no longer pursuing the development of this intellectual property. Following study completion, HK, PM, and LO ended their research relationship with Vital Labs, Inc. IE maintains an interest in Vital Labs. The other authors have not reported any financial disclosures.

Multimedia Appendix 1
Library of patient-reported outcome measures for cystic fibrosis and inflammatory bowel disease.
[DOCX File , 29 KB - mhealth_v8i11e11968_app1.docx]

Multimedia Appendix 2
Parent visit engagement and parent care impact surveys.
[DOCX File , 44 KB - mhealth_v8i11e11968_app2.docx]

References

1. Seid M, Margolis PA, Opipari-Arrigan L. Engagement, peer production, and the learning healthcare system. JAMA Pediatr 2014 Mar;168(3):201-202 [FREE Full text] [doi: 10.1001/jamapediatrics.2013.5063] [Medline: 24446048]
2. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. Milbank Q 1996;74(4):511-544. [Medline: 8941260]
3. Margolis PA, Peterson LE, Seid M. Collaborative chronic care networks (C3Ns) to transform chronic illness care. Pediatrics 2013 Jun;131(Suppl 4):S219-S223 [FREE Full text] [doi: 10.1542/peds.2012-3786J] [Medline: 23729764]
4. Spiegel B. 2015 American journal of gastroenterology lecture: how digital health will transform gastroenterology. Am J Gastroenterol 2016 May;111(5):624-630. [doi: 10.1038/ajg.2016.68] [Medline: 27045930]
5. Finkelstein J, Knight A, Marinopoulos S, Gibbons MC, Berger Z, Aboumatar H, et al. Enabling patient-centered care through health information technology. Evid Rep Technol Assess (Full Rep) 2012 Jun(206):1-1531. [Medline: 24422882]
6. 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]
7. Kvedar J, Coye MJ, Everett W. Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff (Millwood) 2014 Feb;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992] [Medline: 24493760]
8. L'Engle KL, Mangone ER, Parcesepe AM, Agarwal S, Ippoliti NB. Mobile Phone Interventions for Adolescent Sexual and Reproductive Health: A Systematic Review. Pediatrics 2016 Sep;138(3):e20160884. [doi: 10.1542/peds.2016-0884] [Medline: 27553221]
9. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3 [FREE Full text] [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]

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. Bigna JJ, Noubiap JJ, Kouanfack C, Plottel CS, Koulla-Shiro S. Effect of mobile phone reminders on follow-up medical care of children exposed to or infected with HIV in Cameroon (MORE CARE): a multicentre, single-blind, factorial, randomised controlled trial. Lancet Infect Dis 2014 Jul;14(7):600-608. [doi: 10.1016/S1473-3099(14)70741-8] [Medline: 24932893]
12. Strandbygaard U, Thomsen SF, Backer V. A daily SMS reminder increases adherence to asthma treatment: a three-month follow-up study. Respir Med 2010 Feb;104(2):166-171 [FREE Full text] [doi: 10.1016/j.rmed.2009.10.003] [Medline: 19854632]
13. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Diabet Med 2006 Dec;23(12):1332-1338. [doi: 10.1111/j.1464-5491.2006.01989.x] [Medline: 17116184]
14. Khonsari S, Subramanian P, Chinna K, Latif LA, Ling LW, Gholami O. Effect of a reminder system using an automated short message service on medication adherence following acute coronary syndrome. Eur J Cardiovasc Nurs 2015 Apr;14(2):170-179. [doi: 10.1177/1474515114521910] [Medline: 24491349]
15. Carlsen K, Jakobsen C, Houen G, Kallemose T, Paerregaard A, Riis LB, et al. Self-managed ehealth disease monitoring in children and adolescents with inflammatory bowel disease: a randomized controlled trial. Inflamm Bowel Dis 2017 Mar;23(3):357-365. [doi: 10.1097/MIB.0000000000001026] [Medline: 28221247]
16. Fiks AG, Mayne SL, Karavite DJ, Suh A, O'Hara R, Localio AR, et al. Parent-reported outcomes of a shared decision-making portal in asthma: a practice-based RCT. Pediatrics 2015 Apr;135(4):e965-e973 [FREE Full text] [doi: 10.1542/peds.2014-3167] [Medline: 25755233]
17. Kew KM, Cates CJ. Home telemonitoring and remote feedback between clinic visits for asthma. Cochrane Database Syst Rev 2016 Aug 3(8):CD011714 [FREE Full text] [doi: 10.1002/14651858.CD011714.pub2] [Medline: 27486836]
18. Collins F. How to fulfill the true promise of 'mHealth': mobile devices have the potential to become powerful medical tools. Sci Am 2012 Jul;307(1):16. [doi: 10.1038/scientificamerican0712-16] [Medline: 22779258]
19. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
20. Noah B, Keller MS, Mosadeghi S, Stein L, Johl S, Delshad S, et al. Impact of remote patient monitoring on clinical outcomes: an updated meta-analysis of randomized controlled trials. NPJ Digit Med 2018;1:20172 [FREE Full text] [doi: 10.1038/s41746-017-0002-4] [Medline: 31304346]
21. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med 2013;10(2):e1001382 [FREE Full text] [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]
22. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 4;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
23. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR Mhealth Uhealth 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]
24. Gee PM, Greenwood DA, Paterniti DA, Ward D, Miller LM. The ehealth enhanced chronic care model: a theory derivation approach. J Med Internet Res 2015 Apr 1;17(4):e86 [FREE Full text] [doi: 10.2196/jmir.4067] [Medline: 25842005]
25. Jimison H, Gorman P, Woods S, Nygren P, Walker M, Norris S, et al. Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. Evid Rep Technol Assess (Full Rep) 2008 Nov(175):1-1422. [Medline: 19408968]
26. Miyamoto SW, Henderson S, Young HM, Pande A, Han JJ. Tracking health data is not enough: a qualitative exploration of the role of healthcare partnerships and mhealth technology to promote physical activity and to sustain behavior change. JMIR Mhealth Uhealth 2016 Jan 20;4(1):e5 [FREE Full text] [doi: 10.2196/mhealth.4814] [Medline: 26792225]
27. Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. BMC Health Serv Res 2013 Jun 11;13:211 [FREE Full text] [doi: 10.1186/1472-6963-13-211] [Medline: 23758898]
28. Young H, Miyamoto S, Ward D, Dharmar M, Tang-Feldman Y, Berglund L. Sustained effects of a nurse coaching intervention via telehealth to improve health behavior change in diabetes. Telemed J E Health 2014 Sep;20(9):828-834 [FREE Full text] [doi: 10.1089/tmj.2013.0326] [Medline: 25061688]
29. Kaplan HC, Thakkar SN, Burns L, Chini B, Dykes DM, McPhail GL, et al. Protocol of a pilot study of technology-enabled coproduction in pediatric chronic illness care. JMIR Res Protoc 2017 Apr 28;6(4):e71 [FREE Full text] [doi: 10.2196/resprot.7074] [Medline: 28455274]
30. Orchestra Video. YouTube. URL: https://youtu.be/58pkHqLu6cI [accessed 2020-10-13]
31. Lai J, Stucky BD, Thissen D, Varni JW, DeWitt EM, Irwin DE, et al. Development and psychometric properties of the PROMIS(®) pediatric fatigue item banks. Qual Life Res 2013 Nov;22(9):2417-2427 [FREE Full text] [doi: 10.1007/s11136-013-0357-1] [Medline: 23378106]

32. Varni JW, Stucky BD, Thissen D, Dewitt EM, Irwin DE, Lai J, et al. PROMIS pediatric pain interference scale: an item response theory analysis of the pediatric pain item bank. J Pain 2010 Nov;11(11):1109-1119 [FREE Full text] [doi: 10.1016/j.jpain.2010.02.005] [Medline: 20627819]
33. Provost LP, Murray SK. The Health Care Data Guide: Learning from Data for Improvement. San Francisco, CA: Jossey-Bass; 2011.
34. Crandall WV, Margolis PA, Kappelman MD, King EC, Pratt JM, Boyle BM, ImproveCareNow Collaborative. Improved outcomes in a quality improvement collaborative for pediatric inflammatory bowel disease. Pediatrics 2012 Apr;129(4):e1030-e1041 [FREE Full text] [doi: 10.1542/peds.2011-1700] [Medline: 22412030]
35. Szczesniak R, Heltshe SL, Stanojevic S, Mayer-Hamblett N. Use of FEV in cystic fibrosis epidemiologic studies and clinical trials: a statistical perspective for the clinical researcher. J Cyst Fibros 2017 May;16(3):318-326 [FREE Full text] [doi: 10.1016/j.jcf.2017.01.002] [Medline: 28117136]
36. Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. Eff Clin Pract 2001;4(6):256-262. [Medline: 11769298]
37. Ritter PL, Lorig K. The English and Spanish self-efficacy to manage chronic disease scale measures were validated using multiple studies. J Clin Epidemiol 2014 Nov;67(11):1265-1273. [doi: 10.1016/j.jclinepi.2014.06.009] [Medline: 25091546]
38. Hsu WC, Lau KH, Huang R, Ghiloni S, Le H, Gilroy S, et al. Utilization of a cloud-based diabetes management program for insulin initiation and titration enables collaborative decision making between healthcare providers and patients. Diabetes Technol Ther 2016 Feb;18(2):59-67 [FREE Full text] [doi: 10.1089/dia.2015.0160] [Medline: 26645932]
39. Gittell JH, Fairfield KM, Bierbaum B, Head W, Jackson R, Kelly M, et al. Impact of relational coordination on quality of care, postoperative pain and functioning, and length of stay: a nine-hospital study of surgical patients. Med Care 2000 Aug;38(8):807-819. [doi: 10.1097/00005650-200008000-00005] [Medline: 10929993]
40. Cramm JM, Nieboer AP. Relational coordination promotes quality of chronic care delivery in Dutch disease-management programs. Health Care Manage Rev 2012;37(4):301-309. [doi: 10.1097/HMR.0b013e3182355ea4] [Medline: 22138737]

Abbreviations

CCH: Cincinnati Children's Hospital
CF: cystic fibrosis
FEV1: forced expiratory volume in 1 second
IBD: inflammatory bowel disease
mHealth: mobile health
PGA: Physician Global Assessment
PRO: patient-reported outcome
PVP: previsit plan

Edited by G Eysenbach; submitted 22.08.18; peer-reviewed by M Ferrer, T Poder, C Maheu; comments to author 07.01.19; revised version received 08.06.20; accepted 01.09.20; published 26.11.20.

Please cite as:
Opipari-Arrigan L, Dykes DMH, Saeed SA, Thakkar S, Burns L, Chini BA, McPhail GL, Eslick I, Margolis PA, Kaplan HC
Technology-Enabled Health Care Collaboration in Pediatric Chronic Illness: Pre-Post Interventional Study for Feasibility, Acceptability, and Clinical Impact of an Electronic Health Record–Linked Platform for Patient-Clinician Partnership
JMIR Mhealth Uhealth 2020;8(11):e11968
URL: https://mhealth.jmir.org/2020/11/e11968
doi:10.2196/11968
PMID:33242014

Original Paper

Effectiveness of Internet-Based Multicomponent Interventions for Patients and Health Care Professionals to Improve Clinical Outcomes in Type 2 Diabetes Evaluated Through the INDICA Study: Multiarm Cluster Randomized Controlled Trial

Yolanda Ramallo-Fariña[1,2,3,4], MSc; Miguel Angel García-Bello[1], MSc; Lidia García-Pérez[1,2,3,4], MSc; Mauro Boronat[5,6], MD, PhD; Ana M Wägner[5,6], MSc, PhD; Leticia Rodríguez-Rodríguez[1,3], BSc; Pedro de Pablos-Velasco[6,7], MD, PhD; Ignacio Llorente Gómez de Segura[8], MD, PhD; Himar González- Pacheco[1,3], MSc; Montserrat Carmona Rodríguez[2,9], PhD; Pedro Serrano-Aguilar[2,3,4], MD, PhD; INDICA Team[10]

[1]Canary Islands Health Research Institute Foundation (FIISC), Tenerife, Spain

[2]Research Network on Health Services in Chronic Diseases (REDISSEC), Madrid, Spain

[3]Evaluation Unit (SESCS), Canary Islands Health Service (SCS), Tenerife, Spain

[4]Center for Biomedical Research of the Canary Islands (CIBICAN), Tenerife, Spain

[5]Department of Endocrinology and Nutrition, Insular University Hospital, Las Palmas de Gran Canaria, Spain

[6]University Institute for Biomedical and Health Research (IUIBS), University of Las Palmas de Gran Canaria (ULPG), Las Palmas de Gran Canaria, Spain

[7]Department of Endocrinology and Nutrition, Doctor Negrín University Hospital, Las Palmas de Gran Canaria, Spain

[8]Department of Endocrinology and Nutrition, Nuestra Señora de la Candelaria University Hospital, Santa Cruz de Tenerife, Spain

[9]Health Technology Assessment Agency, Instituto de Salud Carlos III, Madrid, Spain

[10]See Author´s Contributions Section, Santa Cruz de Tenerife, Spain

Corresponding Author:
Yolanda Ramallo-Fariña, MSc
Canary Islands Health Research Institute Foundation (FIISC)
Camino La Candelaria 44
Tenerife, 38109
Spain
Phone: 34 +34 922 478266
Email: yramfar@sescs.es

Abstract

Background: Type 2 diabetes mellitus (T2DM) is a chronic disease in which health outcomes are related to decision making by patients and health care professionals.

Objective: This study aims to assess the effectiveness of internet-based multicomponent interventions to support decision making of all actors involved in the care of patients with T2DM in primary care.

Methods: The INDICA study is an open, community-based, multicenter trial with random allocation to usual care or the intervention for patients, the intervention for health care professionals in primary care, or the combined intervention for both. In the intervention for patients, participants received an educational group program and were monitored and supported by logs, a web-based platform, and automated SMS. Those in the intervention for professionals also received an educational program, a decision support tool embedded in the electronic clinical record, and periodic feedback about patients' results. A total of 2334 people with T2DM, regardless of glycated hemoglobin (HbA_{1c}) levels and without diabetes-related complications, were included. The primary end point was change in HbA_{1c} level. The main analysis was performed using multilevel mixed models.

Results: For the overall sample, the intervention for patients attained a significant mean reduction in HbA_{1c} levels of 0.27 (95% CI 0.45 to 0.10) at month 3 and 0.26 (95% CI 0.44 to 0.08) at month 6 compared with usual care, which remained marginally significant at month 12. A clinically relevant reduction in HbA_{1c} level was observed in 35.6% (191/537) of patients in the intervention for patients and 26.0% (152/586) of those in usual care at month 12 (P=.006). In the combined intervention, HbA_{1c}

reduction was significant until month 18 (181/557, 32.6% vs 140/586, 23.9%; *P*=.009). Considering the subgroup of patients uncontrolled at baseline, all interventions produced significant reductions in HbA_{1c} levels across the entire study period: 0.49 (95% CI 0.70 to 0.27) for the intervention for patients, 0.35 (95% CI 0.59 to 0.14) for the intervention for professionals, and 0.35 (95% CI 0.57 to 0.13) for the combined intervention. Differences in HbA_{1c} for the area under the curve considering the entire period were significant for the intervention for patients and the combined intervention compared with usual care (*P*=.03 for both). Compared with usual care, the intervention for professionals and the combined intervention had significant longer-term reductions in systolic and diastolic blood pressure.

Conclusions: In uncontrolled patients, the intervention for patients at baseline provided clinically relevant and significant longer-term reductions of HbA_{1c} levels. The intervention for professionals and combined intervention also improved the cardiovascular risk profile of patients.

Trial Registration: ClinicalTrials.gov NCT01657227; https://clinicaltrials.gov/ct2/show/NCT01657227

(JMIR Mhealth Uhealth 2020;8(11):e18922) doi:10.2196/18922

KEYWORDS

behavior modification; primary care; type 2 diabetes mellitus; patients adherence; eHealth

Introduction

Background

Type 2 diabetes mellitus (T2DM) is a chronic condition in which long-term health outcomes are related to patients' adherence to lifestyle modifications and pharmacologic treatments. Other stakeholders, such as relatives and primary health care professionals, are also involved in guiding patients' decisions.

Although the prevalence of T2DM in the Canary Islands is slightly higher than the average in Spain [1], the incidence of chronic diabetes-related complications [2,3] and mortality [4] is much greater. This occurs despite a continuous increase in diabetes-related public expenditure [5].

Regardless of the widespread availability of evidence-based clinical practice guidelines (CPGs) to care for T2DM, patients' access to effective educational interventions [6] and adherence to self-management activities remains limited internationally [7].

To address these unmet needs, many publications have reported on the effectiveness of using information and communications technology (ICT) applications to support decision making by patients and professionals [8-12], reporting favorable short-term effects on blood glucose control [11,12]. The effectiveness of other biological, cognitive, behavioral, or emotional outcome measures remains controversial [11]. Few large randomized controlled trials (RCTs) have assessed the long-term effectiveness of multicomponent ICT-based interventions, not only for patients but also for all stakeholders involved in diabetes management.

Objectives

The INDICA study is a cluster RCT conducted in the Canary Islands that assesses the effectiveness and cost-effectiveness of multicomponent interventions to support decision making for the main actors involved in the management of T2DM (patients, relatives, and primary health care professionals) in a large number of primary health care practices (PHCPs) [13]. We hypothesized that combining conventional educational activities with different ICT-based decision support tools would efficiently improve health outcomes in patients with T2DM. The main purpose of this study is to evaluate the long-term clinical effectiveness (24 months) of these multicomponent interventions compared with usual care on glycated hemoglobin (HbA_{1c}).

Methods

Study Design

The INDICA study is an open, community-based pragmatic, multicenter, clinical controlled trial with random allocation by clusters to usual care or to one of the following 3 interventions of knowledge transfer and behavior modification:

- Group 1 included interventions for patients and a family member (intervention for patients)
- Group 2 included interventions for health care professionals (physicians and nurses) at primary care (intervention for professionals)
- Group 3 combined the interventions for patients and professionals (combined intervention)

In the usual care or control group, neither patients or families nor physicians or nurses received any additional educational or supporting activities beyond the usual activities provided by the PHCP. The full study protocol has been reported elsewhere [13].

Study Participants

The INDICA study included patients with T2DM aged between 18 and 65 years, diagnosed at least 1 year before study enrollment, without diabetes-related complications, and who regularly used a mobile phone (Textbox 1 provides more details).

Textbox 1. Patients' inclusion and exclusion criteria.

Patient inclusion criteria:

- Patients with type 2 diabetes mellitus diagnosed at least 1 year before study enrollment
- Aged between 18 and 65 years
- Formal consent to participate in the study
- Regular usage of mobile phone

Patient exclusion criteria:

- Chronic kidney disease ≥ stage 3b, as defined by the National Kidney Foundation's Kidney Disease Outcomes and Quality Improvement Initiative, urinary albumin to creatinine ratio ≥ 300 mg/g, or urinary protein excretion ≥ 300 mg/24 hours
- Acute coronary syndrome (documented angina or myocardial infarction) or stroke in the last 6 months or class III or IV heart failure, according to the New York Heart Association
- Proliferative diabetic retinopathy or clinically significant diabetic macular edema requiring previous treatment with retinal photocoagulation, vitrectomy, or intravitreal injections of antivascular endothelial growth factor or triamcinolone acetonide 6 months before study inclusion
- Uncorrected severe hearing or visual impairment or corrected visual acuity ≤ 20/40 by any cause
- Diabetic foot with ulcers ≥ 2 according to the Wagner scale
- Liver cirrhosis
- Cancer, unless disease free 5 years after diagnosis
- Other terminal illnesses
- Intellectual retardation, dementia, and psychotic diseases
- Active substance abuse, alcohol, or drugs (must be sober for 1 year)
- Pregnancy
- Insufficient (Spanish) language skills
- Physical disability limiting participation in group education activities
- Concurrent participation in another clinical trial or any other investigational study

The family care unit (FCU) in each PHCP, comprising a family physician and a nurse responsible for the same set of patients, was the unit of recruitment. FCUs either planning or awaiting placement changes among PHCP in the first 6 months after project initiation were excluded.

All PHCPs included had to have at least eight FCUs and the availability of appropriate places to provide educational group sessions.

Setting and Recruitment

PHCPs were recruited in 4 Canary Islands (Tenerife, Gran Canaria, Lanzarote, and La Palma). FCUs were randomly selected from all consenting FCUs at each PHCP. The electronic clinical records (ECRs) of all potentially eligible patients in all selected FCUs were screened to verify inclusion and exclusion criteria. Finally, eligible patients were randomly selected per FCU.

Random Assignment

Randomization was performed at different levels. First, 3 different strata were created according to the geographical areas in the more populated islands (Tenerife and Gran Canaria). Second, 4 PHCP (clusters) were randomly allocated to every geographical stratum, and block permutation was used to assign PHCPs to the study arms (in total 12 PHCPs for each island), with PHCP as the sampling unit. La Palma and Lanzarote (less populated islands) were geographically divided into 4 zones with only 1 eligible PHCP available in each zone, which was randomly assigned to one of the study arms. On every island, all arms were equally distributed. A total of 6 FCUs were randomly selected from all those consenting to participate in each PHCP. Furthermore, 15 patients were randomly selected from all patients fulfilling the inclusion criteria and consenting to participate in each FCU. Exceptionally, more than 6 FCUs or more than 15 patients per FCU were selected to recruit 90 patients at every PHCP.

FCU and patient randomizations were performed by simple generation from a list of random numbers.

Cluster allocation avoids contamination bias among participants, also facilitating logistics in group interventions.

Interventions

Patient Interventions

Patients recruited to the intervention for patients and combined intervention groups received a complex intervention of knowledge transfer and behavior modification, informed by conceptual frameworks of behavioral change [14]. Key determinants of behavior change suggested by Michie et al [14] were considered for intervention design and implementation, including social and professional role and identity, knowledge, skills, beliefs about capabilities, beliefs about consequences,

motivation and goals, memory, attention and decision processes, environmental context and resources, social influences, emotion, and action planning. Linked to these construct domains, interventions included all techniques judged as effective by the same authors [14], combining (1) a conventional group educational program with a set of 8 quarterly 3-hour group sessions; (2) monitoring of physical activity, diet, drug adherence, mood, blood pressure, and blood glucose readings by daily usage of paper workbooks, complemented by weekly access to a website to download paper workbook data (Multimedia Appendix 1); and (3) continuous personalized feedback by semiautomated mobile phone messages based on the results from the website.

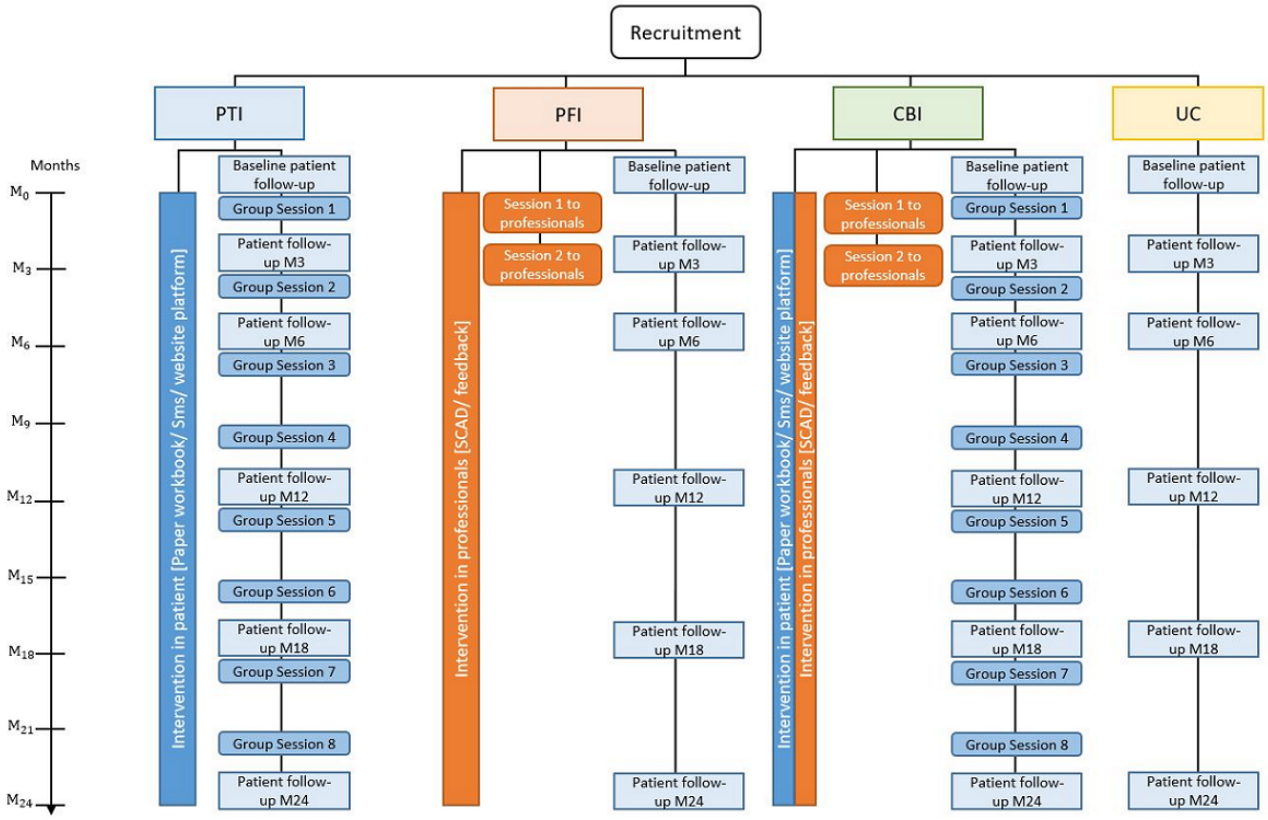
Interventions for Primary Care Professionals

Primary care professionals recruited to the intervention for professionals and combined intervention groups received a complex intervention of knowledge transfer and decision support, partially addressing the determinants of behavior change suggested by Michie et al [14] for its design and implementation, including only techniques to improve skills, environmental changes, prompts and cues by means of electronic clinical guidelines linked to the ECR, processes for encouraging and supporting doctors and nurses, persuasive communication, and periodic feedback on outcomes compared with other colleagues. The interventions combined (1) an educational and interactive group program of 2 sessions to update clinical management and promote patient-centered care; (2) an automated decision aid tool based on a CPG for T2DM, embedded into the ECR (Multimedia Appendix 2); and (3) monthly computerized graphic feedback, displaying a set of processes and outcome indicators for all patients with T2DM of the corresponding FCU.

To maintain the fidelity of interventions, a manual was developed for each intervention. Furthermore, all group sessions were recorded and reviewed.

Both interventions were applied during the 2 years of follow-up (Figure 1).

Figure 1. Arm's intervention timeline and follow-up points.



Duration of Fieldwork

Fieldwork took place between February 2013 and October 2016. The first year was devoted to the recruitment of patients and health care providers and the following 2 years to the intervention and follow-up. As interventions were maintained over time, the intervention and follow-up periods overlapped.

Outcomes

Primary End Point

The primary outcome was the mean change in HbA_{1c} levels from baseline to 24 months of follow-up. HbA_{1c} was also measured at 3, 6, 12, and 18 months. We considered a change in HbA_{1c} of 0.4% as clinically significant [15], just between the thresholds of 0.3% reported by National Institute for Health and Care Excellence [16] and 0.5% by the United Kingdom Prospective Diabetes Study [17].

Secondary End Points

BMI, weight, waist circumference, waist-to-hip ratio, systolic blood pressure (SBP), and diastolic blood pressure (DBP) were also assessed at baseline and after 3, 6, 12, 18, and 24 months. Total, high-density lipoprotein (HDL), and low-density lipoprotein (LDL) cholesterol, triglycerides, and fasting serum glucose were assessed at baseline and after 6, 12, and 24 months. Serum creatinine and glomerular filtration rate were measured at baseline and at 12 and 24 months. Demographic data and disease history were recorded at baseline. Health status and current medications were also recorded at each follow-up.

Statistical Analysis

The main analysis for primary and secondary end points were multilevel mixed models including the baseline value of the dependent variable and the time elapsed since diagnosis (in years) as covariates. The null hypothesis for each end point is that the mean change with regard to the usual care arm and the interactions between each arm and time (follow-up) are the same across arms and equal to zero. The alternative hypothesis is that the changes are not equal to zero. First-level variables are those corresponding to each measurement along follow-up (repeated time measurements), the second level includes patients' variables, and third-level variables correspond to PHCPs. The mean change was estimated at the observation level. The effect that identifies the intervention arm was considered fixed for the PHCPs, whereas the intercept was considered random. The model also included an interaction term between arm and month, allowing for differences in the intervention effect between follow-up assessments [18]. In addition, to summarize the global treatment effect throughout the whole study period, differences were also calculated for the area under the curve (AUC) of HbA_{1c} and other continuous variables between the different interventions and the usual care group. Furthermore, we examined whether the most intensive intervention, the combined intervention group (intervention for patients plus intervention for professionals), was better than the intervention for patients and intervention for professionals groups on their own.

The adjusted estimated mean was calculated for each moment of follow-up compared with baseline, and its significance was calculated using the model already set out.

A post hoc analysis was performed for the primary end point, HbA_{1c}, considering the patient subsample with baseline HbA_{1c} higher than 7%.

To accommodate missing values in the effect analyses, the multiple imputation procedure in Stata 15.0 software (Stata Corporation) was used [19], with results based on 100 imputed data sets. This procedure saves cases for the analysis and can be considered an intention-to-treat analysis. Analysis under multiple imputation is valid for randomly missed data [20]. The model of imputation used and further details on data analysis are outlined in Multimedia Appendix 3. A threshold of .05 was used to define the statistical significance of those tests.

Sample Size Calculation

We estimated the sample size requirement of 448 patients per study arm to detect an absolute difference in HbA_{1c} of 0.4%, assuming a common standard deviation of 1.4% [15], a two-tailed power of 90%, an alpha of .05, and an adjustment for clustering of patients within the FCU by the design effect [21], 15 patients per FCU, and an intraclass correlation coefficient (ICC) of 0.01 (interquartile range 0-0.032) [22]. The intraclass correlation within PHCPs was insignificant as they are formed of several FCUs sharing administrative management and some additional services whose potential effects were already controlled by means of the stratification. Despite this consideration, the sample size was increased by an additional 30% to accommodate for expected losses to follow-up and to warrant the presence of each study arm in all islands. Hence, we aimed to obtain a total sample size of 2330.

Ethics Approval and Consent to Participate

All participants provided written informed consent. The scientific and ethics committees of both the University Hospital of Canarias and the University Hospital Nuestra Señora de la Candelaria approved the study protocol. The study was performed in accordance with Good Clinical Practice standards, applicable local regulatory requirements, and the recommendations of the Declaration of Helsinki.

Results

Study Participants

A total of 32 PHCPs with a mean of 6.6 (SD 0.9) FCUs were included (211 professionals), with 8 PHCPs allocated to each of the 4 study arms. Every PHCP enrolled a mean of 72.9 (SD 14.1) patients (12 patients per FCU), totaling 2334 patients. Figure 2 shows the flowchart for the patients taking part.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.

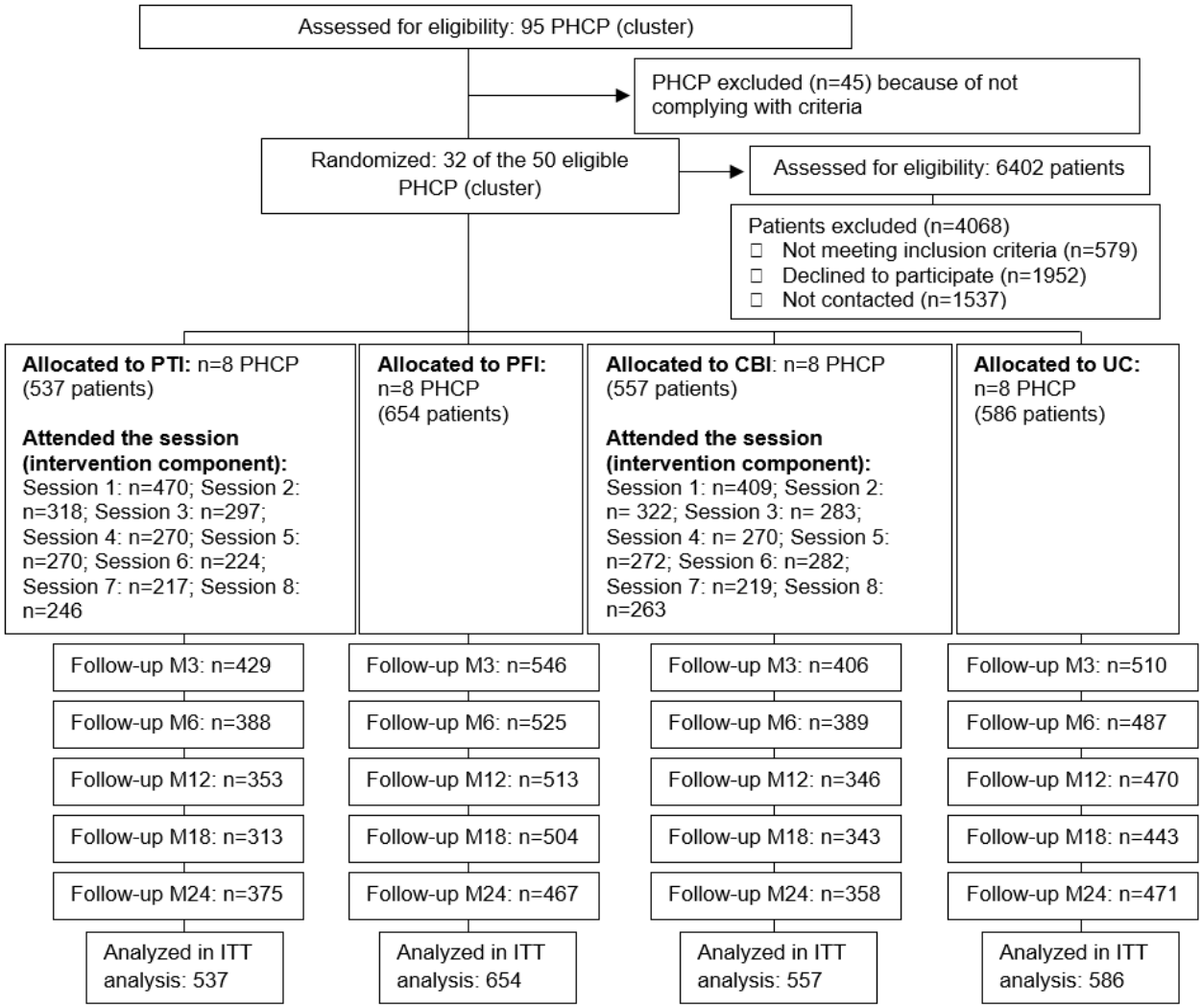


Table 1 shows the patients' baseline characteristics according to the intervention assignment. The mean age of the whole population was 55.7 (SD 7.1) years, with 51.9% (1212/2334) being women. The mean basal HbA_{1c} value was 7.3% (SD 1.5). Overall, 53.4% (1246/2334) of patients had HbA_{1c} levels within the accepted therapeutic goal (≤7%). There were no statistically significant differences among the groups in terms of their baseline characteristics.

Table 1. Baseline characteristics of patients.

Characteristics	PTI[a] (n=537)	PFI[b] (n=654)	CBI[c] (n=557)	UC[d] (n=586)
Age (years), mean (SD)	55.9 (7.0)	56.2 (7.0)	55.5 (7.1)	55.2 (7.3)
Gender (male), n (%)	284 (52.9)	288 (44.0)	264 (47.4)	286 (48.8)
Smoking status, n (%)				
Current smokers	114 (21.2)	156 (23.9)	109 (19.6)	145 (24.7)
Former smokers	223 (41.5)	280 (42.8)	225 (40.4)	240 (41.0)
Nonsmoker	200 (37.2)	218 (33.3)	223 (40.0)	201 (34.3)
Education, n (%)				
Primary or less	323 (60.2)	409 (62.5)	347 (62.3)	379 (64.7)
High school	159 (29.6)	176 (26.9)	147 (26.4)	157 (26.8)
Bachelor's degree or higher	55 (10.2)	69 (10.6)	63 (11.3)	50 (8.5)
Income per person in the household per month, n (%)				
<€250 (US $325)	118 (21.9)	139 (21.2)	121 (21.7)	146 (24.9)
€250-€499 (US $325-$649)	229 (42.7)	323 (49.4)	272 (48.8)	264 (45.1)
€500-€649 (US $650-$844)	86 (16.0)	99 (15.2)	122 (14.2)	96 (16.3)
>€750 (US $975)	104 (19.4)	93 (14.2)	64 (15.3)	80 (13.7)
BMI categories, n (%)				
Normal or underweight (<25)	52 (9.7)	58 (8.9)	45 (8.1)	44 (7.5)
Preobese (<30)	164 (30.5)	183 (28.0)	181 (32.5)	197 (33.6)
Obese class 1 (<35)	200 (37.2)	227 (34.7)	175 (31.4)	195 (33.3)
Obese class 2 (<40)	77 (14.3)	122 (18.7)	103 (18.5)	99 (16.9)
Obese class 3 or 4 (≥40)	44 (8.2)	64 (9.8)	53 (9.5)	51 (8.7)
BMI (kg/m^2), mean (SD)	31.6 (5.7)	32.4 (6.0)	32.1 (5.8)	32.1 (6.0)
Duration of diabetes (years), mean (SD)	8.4 (6.8)	8.2 (6.1)	8.9 (6.3)	8.6 (6.8)
Diabetes treatment, n (%)				
Only lifestyle	40 (7.5)	60 (9.2)	26 (4.7)	53 (9.0)
Oral	394 (73.4)	445 (68.0)	413 (74.1)	395 (67.4)
Injectable (insulin or GLP-1[e])	12 (2.2)	17 (2.6)	17 (3.1)	25 (4.3)
Oral+injectable	85 (15.8)	114 (17.4)	98 (17.6)	98 (16.7)
Do not know/not answered	6 (1.1)	18 (2.8)	3 (0.5)	15 (2.6)
HbA_{1c}[f] categories, n (%)				
<7%	258 (48.0)	351 (53.7)	241 (43.3)	304 (51.9)
7.0%-8.0%	146 (27.2)	165 (25.2)	165 (29.6)	141 (24.1)
8.1%-9.0%	66 (12.3)	75 (11.5)	82 (14.7)	67 (11.4)
>9.0%	67 (12.5)	63 (9.6)	69 (12.4)	74 (12.6)
HbA_{1c} (%), mean (SD)	7.3 (1.5)	7.2 (1.4)	7.4 (1.5)	7.3 (1.5)
Comorbidities, n (%)				
Hypertension	323 (60.3)	434 (66.4)	363 (65.2)	382 (67.1)
Hypercholesterolemia	353 (65.7)	448 (68.0)	349 (62.7)	367 (64.0)
Coronary artery disease	32 (6.0)	39 (6.0)	26 (4.67)	27 (3.9)
Ictus	12 (2.2)	5 (0.8)	13 (2.3)	14 (2.1)
Thyroid gland disorders	68 (12.7)	76 (11.6)	57 (9.7)	57 (11.8)

[a]PTI: intervention only for patients and family members.

[b]PFI: intervention only for health care professionals at primary care.

[c]CBI: combined intervention for patients and professionals.

[d]UC: usual care or control group.

[e]GLP-1: glucagon-like peptide-1.

[f]HbA_{1c}: glycated hemoglobin.

The rate of attendance at educational sessions is also shown in Figure 2. The mean number of sessions attended by patients in the intervention for patients and combined intervention groups was 4.3 (SD 2.7) and 4.2 (SD 2.8), respectively. Overall, 87.5% (470/537) of the patients assigned to the intervention for patients group attended the first of the 8 educational sessions, which decreased to 59.2% (318/537) in the second session and to 45.8% (246/537) in the last session. In the combined intervention group, attendance rates were 73.4% (409/557), 57.8% (322/557), and 47.2% (263/557), respectively. All patients in the intervention groups received SMS during the 2 years of follow-up and had access to the web platform that contained the video recordings of all group sessions in addition to other educational materials. The average number of web-based questionnaires filled in by each patient was 16.3 (SD 29.4) in the intervention for patients group and 9.9 (SD 23.1) in the combined intervention group. These differences were statistically significant (P<.001) at the 2-year follow-up.

Primary End Point: HbA1c

Multimedia Appendix 4 shows the adjusted differences in the mean HbA_{1c} levels at each follow-up evaluation and the adjusted differences in AUCs of HbA_{1c} throughout the whole study for each intervention group, in comparison with the usual care group. Compared with usual care, intervention for patients achieved a significant mean HbA_{1c} reduction of 0.27 (95% CI 0.45 to 0.10) at month 3 and 0.26 (95% CI 0.44 to 0.08) at month 6. Differences between intervention for patients and usual care groups were marginally significant at 12 months (P=.07). There were no statistically significant differences in mean HbA_{1c} levels in the intervention for professionals and combined intervention groups, when compared with the usual care group. With regard to the AUC of HbA_{1c}, the effect of intervention for patients was marginally significant compared with usual care (P=.06), considering all the follow-up sessions.

The mean levels of HbA_{1c} across the study and their adjusted differences with regard to baseline values are shown in Multimedia Appendix 5 by the study arm. Mean HbA_{1c} levels of the intervention for patients group significantly improved during the first 12 months of follow-up, showing a maximal reduction at month 3 (0.35; 95% CI 0.48 to 0.22). The differences gradually diminished over time until they disappeared at months 18 and 24.

At month 3, a clinically relevant reduction in HbA_{1c} (at least 0.4%) was observed in 38.6% (207/537) of participants in the intervention for patients group and only in 20.3% (119/586) of patients with usual care (P<.001; Multimedia Appendix 6). Differences between both groups in the proportion of subjects with a clinically significant decrease in HbA_{1c} remained statistically significant until month 12 (191/537, 35.6% vs 152/586, 26.0%; P=.006) and marginally significant until month 18. The percentage of patients with clinically relevant decrease in HbA_{1c} was also significantly greater in the combined intervention group than in the usual care group at months 3, 6, and 18.

The results of the interventions were also analyzed in the relevant subgroup of uncontrolled patients with baseline HbA_{1c} >7%. As shown in Multimedia Appendix 7, for this subgroup, the differences in the HbA_{1c} reduction between the intervention for patients and usual care groups were statistically significant, favoring the intervention for patients group from months 3 to 12. The differences in HbA_{1c} AUC between the intervention groups and the usual care group considering the entire period were statistically significant for the intervention for patients and combined intervention: 0.26 (95% CI 0.48 to 0.04) and 0.25 (95% CI 0.47 to 0.03), respectively. For the intervention for professionals group, the differences were marginally statistically significant (P=.09).

All interventions led to a significant reduction in HbA_{1c} among subjects with baseline HbA_{1c} levels >7% across the entire study period (Multimedia Appendix 8). The differences at 24 months were 0.49 (95% CI 0.70 to 0.27) for intervention for patients, 0.35 (95% CI 0.59 to 0.14) for intervention for professionals, and 0.35 (95% CI 0.57 to 0.13) for combined intervention (Multimedia Appendix 8). Patients with usual care showed significant decreases in HbA_{1c} at months 12, 18, and 24.

Finally, in the subgroup with baseline HbA_{1c} levels >7%, the proportion of subjects with clinically significant reductions in HbA_{1c} (≥0.4%) was greater in the intervention for patients group than in the usual care group until month 12 (140/263, 53.1% vs 116/269, 43.2%; P=.049). The differences between the combined intervention and the usual care groups were significant at month 3 (Multimedia Appendix 6).

Secondary End Points

Compared with usual care, the intervention for professionals group had significantly lower SBP at months 3 and 18 and the combined intervention group had significantly lower SBP at month 24 (Multimedia Appendix 4). Compared with their respective baseline values, mean SBP fell significantly in all study groups, but the difference was greatest for the combined intervention group at 24 months (7.5 mm Hg; 95% CI 9.8 to 5.2; Multimedia Appendix 5). For DBP, compared with usual care, we found significant reductions at months 3 and 24 for intervention for professionals and at months 12 and 24 for combined intervention (Multimedia Appendix 4). When compared with baseline, all groups improved; the maximum reduction was at 24 months for the combined intervention group, with a fall of 6.7 mm Hg (95% CI 8.2 to 5.3; Multimedia

Appendix 5). The intervention for patients did not lead to a significant decrease in blood pressure compared with usual care (Multimedia Appendix 4).

Comparisons in BMI between the intervention for patients and usual care groups only attained statistically significant differences at month 3. None of the other interventions achieved greater BMI reductions than those observed for usual care (Multimedia Appendix 4). Compared with the baseline values, the mean values of BMI decreased in the intervention for professionals group throughout the follow-up and in the usual care group at months 3 and 24. The intervention for patients group experienced the greatest improvement and showed a statistically significant reduction at month 24: 0.78 kg/m^2 (95% CI 1.0 to 0.6; Multimedia Appendix 5).

Multimedia Appendices 4 and 5 contain detailed biochemical, clinical, and anthropometric data for the whole sample. Multimedia Appendices 7 and 8 contain these data for the subgroup with basal HbA_{1c} >7%.

All 4 groups showed statistically significant improvements in total and LDL cholesterol levels at the end of follow-up. The differences between the intervention and usual care groups were not statistically significant. HDL cholesterol and triglyceride levels did not reveal clinically relevant changes.

We did not detect statistically significant differences in the comparison of intervention for patients and intervention for professionals groups in relation to the most intensive intervention in the combined intervention group regarding the primary or secondary outcomes in the AUC over the follow-up period, except for BMI, which had a difference in area of −0.29 (95% CI −0.57 to 0.01) kg/m^2 in favor of the intervention for patients group.

For most clinical results, ICC values were low in every PHCP. Variance homogeneity was verified and thus reflected a very small effect associated with PHCP for intervention and control groups (similar clinical results among PHCP in every study arm). The ICC at the patient level was broad, accounting for considerable variations among individuals. Considering both ICC values, the results from the INDICA study appear to have good external validity.

Discussion

Principal Findings

The INDICA study assessed the effectiveness of multicomponent interventions to support decision making for the main actors involved in the management of T2DM (patients, relatives, and primary health care professionals) in many PHCPs [13]. We hypothesized that combining conventional educational activities with different ICT-based decision support tools would improve HbA_{1c} at long term (24 months) compared with usual care.

This study revealed that the intervention for patients group achieved a significant but temporary reduction of HbA_{1c}, compared with the usual care group, which lasted for 6 months, with a gradual dilution effect from then onward. Interventions focused on health care professionals and on both patients and health care professionals did not translate into a significant lowering of HbA_{1c}, in comparison with usual care, when evaluated in the whole study population. Even so, more than 30% of the participants belonging to the intervention for patients and combined intervention groups attained statistically and clinically relevant reductions in HbA_{1c} (>0.4%). These percentages were significantly greater than those observed in the control group at 12 months (for the intervention for patients group) and 18 months (for the combined intervention group).

It must be noted that, with the intention of assessing the effectiveness of the intervention for all patients with T2DM, the INDICA study did not limit inclusion of participants by their HbA_{1c} level. Therefore, the study's power to find clinically relevant differences for the main outcome measures could have been insufficient, according to Jackson et al [23], as only 50.6% (1180/2334) of all participants had baseline HbA_{1c} concentrations >7% (mean 7.3%, SD 1.5). Nonetheless, the study's sample size provided statistical power to examine the results of patients with worse metabolic control, allowing the comparison with other studies that limited recruitment to patients with poor metabolic control.

As expected, the magnitude and duration of the intervention effect was greater among patients with baseline HbA_{1c} >7%, mainly for the intervention for patients group, which showed a statistically significant reduction in HbA_{1c}, in comparison with usual care, although the difference disappeared at 18 months. Moreover, considering the differences in the AUC values of HbA_{1c}, our results provide evidence of effectiveness for both the intervention for patients and the combined intervention throughout the study period. These results support previous findings reporting greater effects for interventions on patients with higher baseline HbA_{1c} levels [24,25]. Similarly, the effectiveness of quality improvement strategies exclusively focused on health care providers seems to be beneficial only among patients with HbA_{1c} levels >8% [26].

The Mobile Diabetes Intervention Study (MDIS) published by Quinn et al [27] also reported a higher reduction in HbA_{1c} over 1 year among patients with T2DM (with baseline HbA_{1c}=9.1%) by means of a multicomponent behavioral intervention exclusively for patients, without detecting effects on other relevant outcomes such as blood pressure or lipid levels.

Although MDIS provided evidence of sustained 12-month treatment difference in HbA_{1c}, rather than *regression to the mean*, the INDICA results, for the whole sample, show a progressive effect reduction close to the baseline HbA_{1c} values. Similar to MDIS, the observed reduction in HbA_{1c} in the INDICA subgroup with baseline HbA_{1c}>7% remained stable over the long term. However, evidence of long-term effectiveness of these complex interventions is not well stated yet because of the reduced number of studies providing results at 12 months of follow-up and beyond [28,29].

Several systematic reviews found that interventions based on ICTs led to significant improvements of 4% to 5% in HbA_{1c}

compared with usual care [12,28,30,31], with effect differences according to the type of ICT used (internet, automated SMS, and apps) [11,12,32]. In contrast, smaller effects than those reported in our study for the intervention for patients and combined intervention groups were published for individual and group education among patients with HbA_{1c} levels >8% [33,34].

Beyond the reported effects on HbA_{1c}, we also found an improvement in blood pressure monitoring for patients included in the 2 groups with intervening health professionals. Long-term reductions compared with the baseline were observed in SBP and DBP, with statistically significant differences in relation to usual care. These combined effects on HbA_{1c}, SBP, and DBP, together with the improvement observed for BMI, might contribute to enhanced cardiovascular risk [35,36], suggesting the overall value of these comprehensive approach strategies addressing multiple components and actors involved in T2DM management [37]. Although some outcomes, such as the improvement of blood pressure, might require the involvement of health care providers, others, such as the reduction in HbA_{1c}, will depend largely on the patients' intervention. Thus, our findings provide long-term evidence on the effectiveness of multicomponent interventions to empower patients and support clinical decision making to improve T2DM outcomes beyond that published by Taylor et al [29] in their systematic review for self-management interventions for patients with chronic conditions. The potential expected clinical benefits, associated with the overall metabolic and cardiovascular risk improvement provided by INDICA over 2 years, could be estimated in the longer-term follow-up on both microvascular and macrovascular complications and mortality [38].

Conceptual Frameworks

The assessed interventions were informed by conceptual frameworks of behavioral change [14] and applied to a large and heterogeneous sample of patients, caregivers, and professionals. The INDICA intervention characteristics were planned to increase the validity of the obtained data and the transferability of the interventions assessed. The key determinants of behavior change suggested by Michie et al [14] were considered for the INDICA interventions, with a higher degree of adherence in their design and implementation in the case of interventions for patients than for professionals, which could help explain the magnitude of the effect observed for HbA_{1c} among intervention groups. Furthermore, time constraints, staff turnover, and self-perception of work overload among health professionals limited the possibility of going deeper into the following dimensions: professional role, motivation and goals, social and professional influences, emotions, and action planning. A detailed description of the complex behavior change interventions applied was reported elsewhere [13] to promote replication at other sites. Other potential explanations for the unexpected differences between the intervention for patients and combined intervention groups were the higher attendance rate of patient and family members in the educational group sessions and a significantly higher rate of web questionnaire completion observed in the intervention for patients group. This higher rate of questionnaire completion was key to adjusting the individualized components of SMS messages, providing an extended exposure to web-based educational material. The high turnover among health care professionals in most PHCPs included in the study, as occurs in the real world, could also account for the lesser effect of the intervention for professionals and the combined intervention.

To maximize effectiveness, the INDICA interventions incorporated all the components of a technology-enabled self-management feedback cycle, connecting patients and the research team by using bidirectional communication, analyzing patient-provided behavior and health data, tailoring education, and personalizing feedback according to the eHealth Enhanced Chronic Care Model [24,39,40].

Strengths and Limitations

This study has some limitations. First, it was difficult to obtain a full data set because of the high number of control visits and the duration of follow-up for many patients. Robust imputation techniques [19] were used to minimize the impact of missing data. Second, as previously mentioned, the high turnover among primary care professionals included in the study could explain the smaller than expected impact of the intervention for professionals and the combined intervention. Third, the fact that around 49.4% (1154/2334) of the whole patient sample had baseline HbA_{1c} <7% and only around 23.0% (536/2334) had basal HbA_{1c} levels ≥8% clearly limits the ability of interventions to reduce HbA_{1c}. Fortunately, the available sample size was sufficient to find valid evidence. Fourth, similar to other reported findings [23,26,41], our usual care group was not a proper control group; it was subject to repetitive and intensive follow-up activities, including 6 different follow-up visits over the study to apply all prespecified questionnaires, in addition to clinical and laboratory tests. This intense follow-up activity could act as an intervention in itself, as patients might focus on important topics on which they had to pay attention. Fifth, INDICA interventions were not fully theory-based, making it more difficult to understand as to what works across contexts, populations, and behaviors. Finally, the INDICA study was not designed to test the efficacy of every component of the complex interventions assessed.

The strengths of the INDICA study include the pragmatic character of the trial and its wide sample size; the random assignment by clusters; the engagement, as research subjects, of all actors involved in management decisions; and the follow-up duration. Moreover, all educational group sessions and coaching activities by SMS were recorded to monitor and assess homogeneity, educator fidelity to interventions, and quality delivery. Educational workshops and periodic feedback to health care professionals were equally delivered to all participants in the intervention for professionals and the combined intervention.

The INDICA findings highlight the importance of conducting trials with long follow-up periods and sufficient statistical power to assess interventions of limited expected effect sizes but of high potential efficiency. ICT-supported interventions enable its extended and continuous usage by thousands of people in need to complement and spread interventions beyond the limited

capacity of the health care systems to deliver usual care. We should be careful, however, to generalize the findings of INDICA. Interventions took place through PHCPs and were largely implemented through electronic communications. Health and digital literacy levels of the assessed population might vary with regard to other settings. Moreover, health care professionals were subject to differences in workload, interest and training in ICT used to support patients, access to CPG, and specialist support.

The potential effects of all these factors on the different study arms were minimized by randomization.

Future Research

Future research on the effectiveness of these complex interventions should be complemented by the analysis of patients' self-reported outcomes and intervention cost-effectiveness to fully inform clinical and health policy decision making. The effectiveness of these interventions should also be assessed after longer follow-up periods to allow the measurement of relevant clinical (micro and macrovascular) outcomes, together with the assessment of potential longer-term reinforcement of the most cost-effective interventions in the short term. The use of real-world data will efficiently help to provide this valuable information. Effectiveness and cost-effectiveness assessment according to patients' clinical risk and health literacy levels are also highly relevant. Additional evidence on cost-effectiveness and budget impact analysis is needed to support health policy decision making in cases of limited funding to support all assessed interventions.

Theory-based research on complex interventions to promote behavior change is also needed, rather than theory-inspired research, if we are to achieve a sound scientific basis for the development and reporting of such interventions. Comparative effectiveness assessment among components of complex interventions is also of interest, although it will require additional funding.

Finally, qualitative research is also needed to better understand the relationships between patient and professional characteristics, their engagement, and the observed results.

Conclusions

We found that INDICA interventions improved long-term metabolic control in patients with T2DM with uncontrolled basal HbA_{1c} values compared with the usual care group. We also found moderate but clinically and statistically significant effects on blood pressure reduction, contributing to reduced overall cardiovascular risk. The increasing access to computers, internet, and mobile phones, together with improvements in digital literacy, regardless of social status, sex, and age, make these complex interventions appropriate instruments to improve patient empowerment in the continuous management of their chronic diseases by tailoring interventions to individual needs and extending patient support beyond the limited capacities of conventional office-based care.

Acknowledgments

The authors thank Jason Willis-Lee for copyediting services during the preparation of the final manuscript and Thayli León Plasencia for her help in recruiting patients.

This study received financial support from the Spanish Ministry of Economy, Industry, and Competitiveness (Instituto de Salud Carlos III), grant numbers ADE10/00032 and PI16/00769, cofounded by Fondo Europeo de Desarrollo Regional *Una manera de hacer Europa*. The funders did not participate in the study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication.

Authors' Contributions

The INDICA team included the following members: Abraham Pérez de la Rosa (Canary Islands Health Research Institute Foundation, FIISC), Alicia Pareja Ríos (University Hospital of Canary Island), Andrés Sifre Perello (Molina Orosa Hospital), Ángela Trinidad Gutiérrez Pérez (Primary Care of Gran Canaria), Antonio Cabrera de León (Ntra Sra de la Candelaria University Hospital), Antonio García Quintana (Dr. Negrín University Hospital), Armando Carrillo Domínguez (Insular University Hospital), Bernardo Eusebio Herrera Domínguez (General de La Palma Hospital), Carlos Sedeño Pérez (Primary Care of Tenerife), Carlos Ramírez Álamo (Primary Care of Gran Canaria), Carmen Daranas Aguilar (Canary Islands Health Research Institute Foundation, FIISC), Carolina Guerra Marrero (Canary Islands Health Research Institute Foundation, FIISC), Cecilia Lobos Soto (Insular University Hospital), Cristina Padrón Pérez (Canary Islands Health Research Institute Foundation, FIISC), Dácil Alvarado Martel (Dr. Negrín University Hospital), Daniel Hernández Obregón (Dr. Negrín University Hospital), Dulce N. Hernández Correa (Primary Care of Gran Canaria), Elsa Espinosa Pozuelo (Diabetes Patient´ association of Tenerife), Elsa Florido Mayor (Canary Islands Health Research Institute Foundation, FIISC), Engracia Pinilla Domínguez (Ntra Sra de la Candelaria University Hospital), Fátima Herrera García (University Hospital of Canary Island), Félix Bonilla Aguiar (Dr. José Molina Hospital), Fernando Montón Álvarez (Ntra Sra de la Candelaria University Hospital), Francisco Cabrera López (Insular University Hospital), Gloria Guerra de la Torre (Primary Care of Gran Canaria), Gregorio Muelas Martín (Dr. Negrín University Hospital), Guillermo Monzón Monzón (Primary Care of Gran Canaria), Héctor de la Rosa Merino (Canary Islands Health Research Institute Foundation, FIISC), Ignacio García Puente (Dr. Negrín University Hospital), Isabel García Calcerrada (Ntra Sra de la Candelaria University Hospital), Iván Castilla Rodríguez (Canary Islands Health Research Institute Foundation, FIISC), Jacqueline Álvarez Pérez (Canary Islands Health Research Institute Foundation, FIISC), Jorge Federico Aldunate Page (Insular University Hospital), Jose Antonio García

Dopico (University Hospital of Canary Island), Juan Andrés Báez Hernández (Primary Care of La Palma), Juan José Pérez Valencia (Primary Care of Tenerife), Julia Charlotte Wiebe (Dr. Negrín University Hospital), Lilisbeth Perestelo Pérez (Evaluation Unit, SESCS, Canary Islands Health Service, SCS), Leopoldo Martín Martín (Hospital General de La Palma), Lluis Serra Majem (University Institute for Biomedical and Health Research (IUIBS), ULPGC), Luis Morcillo Herrera (University Hospital of Canary Island), Marcos Estupiñán Ramírez (Canary Islands Health Service, SCS), Margarita Roldán Ruano (Primary Care of Gran Canaria), María del Mar Romero Fernández (Canary Islands Health Research Institute Foundation, FIISC), María Inmaculada González Pérez (Ntra Sra de la Candelaria University Hospital), María Isabel Visuerte Morales (University Hospital of Canary Island), María Pino Afonso Medina (Dr. Negrín University Hospital), Marta Riaño Ruiz (Insular University Hospital), Marta Tejera Santana (Dr. Negrín University Hospital), Mercedes Lorenzo Medina (Dr. Negrín University Hospital), Miguel Juan Mora García (Primary Care of Gran Canaria), Nayra Pérez Delgado (Ntra Sra de la Candelaria University Hospital), Pablo Pedrianez Martín (Dr. Negrín University Hospital), Pilar Peláez Alba (La Laguna University), Rafael Valcárcel (Primary Care of Tenerife), Remedios Castro Sánchez (Primary Care of Gran Canaria), Rodrigo Abreu González (Ntra Sra de la Candelaria University Hospital), Rosa Borges Trujillo (Dr. Negrín University Hospital), Salvador Acosta González (Ntra Sra de la Candelaria University Hospital), Sybille Kaiser Girardot (Primary Care of Tenerife), Víctor Lorenzo Sellarés (University Hospital of Canary Island).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of patients' website.
[DOC File , 779 KB - mhealth_v8i11e18922_app1.doc]

Multimedia Appendix 2
Representative screenshots of automated decision aid tool embedded into the electronic clinical record.
[DOC File , 491 KB - mhealth_v8i11e18922_app2.doc]

Multimedia Appendix 3
Multiple imputation model.
[DOC File , 525 KB - mhealth_v8i11e18922_app3.doc]

Multimedia Appendix 4
Adjusted difference in means and area under the curve of each group compared with the usual care group for the whole sample.
[DOC File , 176 KB - mhealth_v8i11e18922_app4.doc]

Multimedia Appendix 5
Adjusted means for each group and intragroup differences compared with the baseline measurement for the whole sample.
[DOC File , 201 KB - mhealth_v8i11e18922_app5.doc]

Multimedia Appendix 6
Patients with clinically relevant changes in glycated hemoglobin and comparison with the usual care group.
[DOC File , 57 KB - mhealth_v8i11e18922_app6.doc]

Multimedia Appendix 7
Adjusted difference in means and area under the curve of each group compared with the usual care group for patients with a baseline glycated hemoglobin level >7%.
[DOC File , 143 KB - mhealth_v8i11e18922_app7.doc]

Multimedia Appendix 8
Adjusted means for each group and intragroup differences compared with the baseline measurement for patients with a baseline glycated hemoglobin level >7 %.
[DOCX File , 61 KB - mhealth_v8i11e18922_app8.docx]

Multimedia Appendix 9
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1620 KB - mhealth_v8i11e18922_app9.pdf]

References

1. Cabrera de León A, Rodríguez Pérez MC, Almeida González D, Domínguez Coello S, Aguirre Jaime A, Brito Díaz B, grupo CDC. [Presentation of the 'CDC de Canarias' cohort: objectives, design and preliminary results]. Rev Esp Salud Publica 2008;82(5):519-534 [FREE Full text] [doi: 10.1590/s1135-57272008000500007] [Medline: 19039505]
2. Lorenzo V, Boronat M, Saavedra P, Rufino M, Maceira B, Novoa FJ, et al. Disproportionately high incidence of diabetes-related end-stage renal disease in the Canary Islands. An analysis based on estimated population at risk. Nephrol Dial Transplant 2010 Jul;25(7):2283-2288. [doi: 10.1093/ndt/gfp761] [Medline: 20064954]
3. Aragón-Sánchez J, García-Rojas A, Lázaro-Martínez JL, Quintana-Marrero Y, Maynar-Moliner M, Rabellino M, et al. Epidemiology of diabetes-related lower extremity amputations in Gran Canaria, Canary Islands (Spain). Diabetes Res Clin Pract 2009 Oct;86(1):e6-e8. [doi: 10.1016/j.diabres.2009.06.015] [Medline: 19604593]
4. Registro De Defunciones Según La Causa De Muerte. Instituto Nacional de Estadística, INE. 2015. URL: https://www.ine.es/jaxi/Datos.htm?path=/t15/p417/a2018/l0/&file=02001.px#!tabs-tabla [accessed 2020-10-01]
5. Lopez-Bastida J, Boronat M, Moreno JO, Schurer W. Costs, outcomes and challenges for diabetes care in Spain. Global Health 2013 May 1;9:17 [FREE Full text] [doi: 10.1186/1744-8603-9-17] [Medline: 23635075]
6. Brown-Guion SY, Youngerman SM, Hernandez-Tejada MA, Dismuke CE, Egede LE. Racial/ethnic, regional, and rural/urban differences in receipt of diabetes education. Diabetes Educ 2013;39(3):327-334. [doi: 10.1177/0145721713480002] [Medline: 23482514]
7. García-Pérez LE, Alvarez M, Dilla T, Gil-Guillén V, Orozco-Beltrán D. Adherence to therapies in patients with type 2 diabetes. Diabetes Ther 2013 Dec;4(2):175-194 [FREE Full text] [doi: 10.1007/s13300-013-0034-y] [Medline: 23990497]
8. Main C, Moxham T, Wyatt JC, Kay J, Anderson R, Stein K. Computerised decision support systems in order communication for diagnostic, screening or monitoring test ordering: systematic reviews of the effects and cost-effectiveness of systems. Health Technol Assess 2010 Oct;14(48):1-227 [FREE Full text] [doi: 10.3310/hta14480] [Medline: 21034668]
9. Årsand E, Frøisland DH, Skrøvseth SO, Chomutare T, Tatara N, Hartvigsen G, et al. Mobile health applications to assist patients with diabetes: lessons learned and design implications. J Diabetes Sci Technol 2012 Sep 1;6(5):1197-1206 [FREE Full text] [doi: 10.1177/193229681200600525] [Medline: 23063047]
10. Saffari M, Ghanizadeh G, Koenig HG. Health education via mobile text messaging for glycemic control in adults with type 2 diabetes: a systematic review and meta-analysis. Prim Care Diabetes 2014 Dec;8(4):275-285. [doi: 10.1016/j.pcd.2014.03.004] [Medline: 24793589]
11. Pal K, Eastwood SV, Michie S, Farmer A, Barnard ML, Peacock R, et al. Computer-based interventions to improve self-management in adults with type 2 diabetes: a systematic review and meta-analysis. Diabetes Care 2014 Jun;37(6):1759-1766. [doi: 10.2337/dc13-1386] [Medline: 24855158]
12. Lee PA, Greenfield G, Pappas Y. The impact of telehealth remote patient monitoring on glycemic control in type 2 diabetes: a systematic review and meta-analysis of systematic reviews of randomised controlled trials. BMC Health Serv Res 2018 Jun 26;18(1):495 [FREE Full text] [doi: 10.1186/s12913-018-3274-8] [Medline: 29940936]
13. Ramallo-Fariña Y, García-Pérez L, Castilla-Rodríguez I, Perestelo-Pérez L, Wägner AM, de Pablos-Velasco P, INDICA team. Effectiveness and cost-effectiveness of knowledge transfer and behavior modification interventions in type 2 diabetes mellitus patients--the INDICA study: a cluster randomized controlled trial. Implement Sci 2015 Apr 9;10:47 [FREE Full text] [doi: 10.1186/s13012-015-0233-1] [Medline: 25880498]
14. 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;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
15. Russell-Jones D, Simpson R, Hylleberg B, Draeger E, Bolinder J. Effects of QD insulin detemir or neutral protamine Hagedorn on blood glucose control in patients with type I diabetes mellitus using a basal-bolus regimen. Clin Ther 2004 May;26(5):724-736. [doi: 10.1016/s0149-2918(04)90072-0] [Medline: 15220016]
16. Type 1 Diabetes in Adults: Diagnosis and Management. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/guidance/ng17/evidence/full-guideline-pdf-435400241 [accessed 2020-10-01]
17. Hayes AJ, Leal J, Gray AM, Holman RR, Clarke PM. UKPDS outcomes model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82. Diabetologia 2013 Sep;56(9):1925-1933. [doi: 10.1007/s00125-013-2940-y] [Medline: 23793713]
18. Finucane MM, Samet JH, Horton NJ. Translational methods in biostatistics: linear mixed effect regression models of alcohol consumption and HIV disease progression over time. Epidemiol Perspect Innov 2007 Sep 19;4:8 [FREE Full text] [doi: 10.1186/1742-5573-4-8] [Medline: 17880699]
19. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011 Feb 20;30(4):377-399. [doi: 10.1002/sim.4067] [Medline: 21225900]
20. Enders C. Applied Missing Data Analysis. New York, USA: The Guilford Press; 2010.
21. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. Br Med J 2012 Sep 4;345:e5661. [doi: 10.1136/bmj.e5661] [Medline: 22951546]

22. Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. J Clin Epidemiol 2004 Aug;57(8):785-794. [doi: 10.1016/j.jclinepi.2003.12.013] [Medline: 15485730]
23. Jackson CL, Bolen S, Brancati FL, Batts-Turner ML, Gary TL. A systematic review of interactive computer-assisted technology in diabetes care. Interactive information technology in diabetes care. J Gen Intern Med 2006 Feb;21(2):105-110. [doi: 10.1111/j.1525-1497.2005.00310.x] [Medline: 16390512]
24. Medical Advisory Secretariat. Behavioural interventions for type 2 diabetes: an evidence-based analysis. Ont Health Technol Assess Ser 2009;9(21):1-45 [FREE Full text] [Medline: 23074526]
25. Peters RM, Lui M, Patel K, Tian L, Javaherian K, Sink E, et al. Improving glycemic control with a standardized text-message and phone-based intervention: a community implementation. JMIR Diabetes 2017 Jul 25;2(2):e15 [FREE Full text] [doi: 10.2196/diabetes.7910] [Medline: 30291063]
26. Tricco AC, Ivers NM, Grimshaw JM, Moher D, Turner L, Galipeau J, et al. Effectiveness of quality improvement strategies on the management of diabetes: a systematic review and meta-analysis. Lancet 2012 Jun 16;379(9833):2252-2261. [doi: 10.1016/S0140-6736(12)60480-2] [Medline: 22683130]
27. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
28. Taylor SJ, Pinnock H, Epiphaniou E, Pearce G, Parke H, Schwappach A, et al. A rapid synthesis of the evidence on interventions supporting self-management for people with long-term conditions: PRISMS – Practical systematic RevIew of Self-Management Support for long-term conditions. Health Serv Deliv Res 2014:-. [doi: 10.3310/hsdr02530] [Medline: 25642548]
29. Clark M. Diabetes self-management education: a review of published studies. Prim Care Diabetes 2008 Sep;2(3):113-120. [doi: 10.1016/j.pcd.2008.04.004] [Medline: 18779034]
30. Wu Y, Yao X, Vespasiani G, Nicolucci A, Dong Y, Kwong J, et al. Mobile app-based interventions to support diabetes self-management: a systematic review of randomized controlled trials to identify functions associated with glycemic efficacy. JMIR Mhealth Uhealth 2017 Mar 14;5(3):e35 [FREE Full text] [doi: 10.2196/mhealth.6522] [Medline: 28292740]
31. Wu X, Guo X, Zhang Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
32. Arambepola C, Ricci-Cabello I, Manikavasagam P, Roberts N, French DP, Farmer A. The impact of automated brief messages promoting lifestyle changes delivered via mobile devices to people with type 2 diabetes: a systematic literature review and meta-analysis of controlled trials. J Med Internet Res 2016 Apr 19;18(4):e86 [FREE Full text] [doi: 10.2196/jmir.5425] [Medline: 27095386]
33. Duke SS, Colagiuri S, Colagiuri R. Individual patient education for people with type 2 diabetes mellitus. Cochrane Database Syst Rev 2009 Jan 21(1):CD005268 [FREE Full text] [doi: 10.1002/14651858.CD005268.pub2] [Medline: 19160249]
34. Deakin T, McShane CE, Cade JE, Williams RD. Group based training for self-management strategies in people with type 2 diabetes mellitus. Cochrane Database Syst Rev 2005 Apr 18(2):CD003417. [doi: 10.1002/14651858.CD003417.pub2] [Medline: 15846663]
35. Cook NR, Cohen J, Hebert PR, Taylor JO, Hennekens CH. Implications of small reductions in diastolic blood pressure for primary prevention. Arch Intern Med 1995 Apr 10;155(7):701-709. [Medline: 7695458]
36. Wing RR, Lang W, Wadden TA, Safford M, Knowler WC, Bertoni AG, Look AHEAD Research Group. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. Diabetes Care 2011 Jul;34(7):1481-1486 [FREE Full text] [doi: 10.2337/dc10-2415] [Medline: 21593294]
37. Mills KT, Obst KM, Shen W, Molina S, Zhang H, He H, et al. Comparative effectiveness of implementation strategies for blood pressure control in hypertensive patients: a systematic review and meta-analysis. Ann Intern Med 2018 Jan 16;168(2):110-120 [FREE Full text] [doi: 10.7326/M17-1805] [Medline: 29277852]
38. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. Br Med J 2000 Aug 12;321(7258):405-412 [FREE Full text] [doi: 10.1136/bmj.321.7258.405] [Medline: 10938048]
39. Gee PM, Greenwood DA, Paterniti DA, Ward D, Miller LM. The ehealth enhanced chronic care model: a theory derivation approach. J Med Internet Res 2015 Apr 1;17(4):e86 [FREE Full text] [doi: 10.2196/jmir.4067] [Medline: 25842005]
40. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-management education and support. J Diabetes Sci Technol 2017 Sep;11(5):1015-1027 [FREE Full text] [doi: 10.1177/1932296817713506] [Medline: 28560898]
41. Anzaldo-Campos MC, Contreras S, Vargas-Ojeda A, Menchaca-Díaz R, Fortmann A, Philis-Tsimikas A. Dulce wireless Tijuana: a randomized control trial evaluating the impact of project dulce and short-term mobile technology on glycemic control in a family medicine clinic in northern Mexico. Diabetes Technol Ther 2016 Apr;18(4):240-251 [FREE Full text] [doi: 10.1089/dia.2015.0283] [Medline: 26914371]

Abbreviations

AUC: area under the curve
CPG: clinical practice guideline
DBP: diastolic blood pressure
ECR: electronic clinical record
FCU: family care unit
HbA1c: glycated hemoglobin
HDL: high-density lipoprotein
ICC: intraclass correlation coefficient
ICT: information and communications technology
LDL: low-density lipoprotein
MDIS: Mobile Diabetes Intervention Study
PHCP: primary health care practice
RCT: randomized controlled trial
SBP: systolic blood pressure
T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 03.04.20; peer-reviewed by K Yakubu, S Lin; comments to author 30.06.20; revised version received 20.08.20; accepted 13.09.20; published 02.11.20.

Please cite as:
Ramallo-Fariña Y, García-Bello MA, García-Pérez L, Boronat M, Wägner AM, Rodríguez-Rodríguez L, de Pablos-Velasco P, Llorente Gómez de Segura I, González- Pacheco H, Carmona Rodríguez M, Serrano-Aguilar P, INDICA Team
Effectiveness of Internet-Based Multicomponent Interventions for Patients and Health Care Professionals to Improve Clinical Outcomes in Type 2 Diabetes Evaluated Through the INDICA Study: Multiarm Cluster Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(11):e18922
URL: https://mhealth.jmir.org/2020/11/e18922
doi:10.2196/18922
PMID:33136059

Original Paper

Smart Data Collection for the Assessment of Treatment Effects in Irritable Bowel Syndrome: Observational Study

Zsa Zsa R M Weerts[1], MD, MSc; Koert G E Heinen[2], MSc; Ad A M Masclee[1], MD, MSc, PhD; Amber B A Quanjel[1], MD, MSc; Bjorn Winkens[3,4], MSc, PhD; Lisa Vork[1], MD, MSc; Paula E L M Rinkens[2], BSc; Daisy M A E Jonkers[1], MSc, PhD; Daniel Keszthelyi[1], MD, MSc, PhD

[1]Division Gastroenterology-Hepatology, Department of Internal Medicine, NUTRIM School for Nutrition and Translational Research in Metabolism, Maastricht University Medical Center+, Maastricht, Netherlands

[2]MEMIC Center for Data and Information Management, Maastricht University, Maastricht, Netherlands

[3]Department of Methodology and Statistics, Maastricht University Medical Center+, Maastricht, Netherlands

[4]Care and Public Health Research Institute, Maastricht University Medical Center+, Maastricht, Netherlands

Corresponding Author:
Zsa Zsa R M Weerts, MD, MSc
Division Gastroenterology-Hepatology, Department of Internal Medicine
NUTRIM School for Nutrition and Translational Research in Metabolism
Maastricht University Medical Center+
Universiteitssingel 50
Maastricht, 6202 AZ
Netherlands
Email: z.weerts@maastrichtuniversity.nl

Related Article:

This is a corrected version. See correction statement: https://mhealth.jmir.org/2021/2/e27998/

Abstract

Background: End-of-day symptom diaries are recommended by drug regulatory authorities to assess treatment response in patients with irritable bowel syndrome. We developed a smartphone app to measure treatment response.

Objective: Because the employment of an app to measure treatment response in irritable bowel syndrome is relatively new, we aimed to explore patients' adherence to diary use and characteristics associated with adherence.

Methods: A smartphone app was developed to serve as a symptom diary. Patients with irritable bowel syndrome (based on Rome IV criteria) were instructed to fill out end-of-day diary questionnaires during an 8-week treatment. Additional online questionnaires assessed demographics, irritable bowel syndrome symptom severity, and psychosocial comorbidities. Adherence rate to the diary was defined as the percentage of days completed out of total days. Adherence to the additional web-based questionnaires was also assessed.

Results: Overall, 189 patients were included (age: mean 34.0 years, SD 13.3 years; female: 147/189, 77.8%; male: 42/189, 22.2%). The mean adherence rate was 87.9% (SD 9.4%). However, adherence to the diary decreased over time (P<.001). No significant association was found between adherence and gender (P=.84), age (P=.22), or education level (lower education level: P=.58, middle education level: P=.46, versus high education level), while higher anxiety scores were associated with lower adherence (P=.03). Adherence to the online questionnaires was also high (>99%). Missing data due to technical issues were limited.

Conclusions: The use of a smartphone app as a symptom diary to assess treatment response resulted in high patient adherence. The data-collection framework described led to standardized data collection with excellent completeness and can be used for future randomized controlled trials. Due to the slight decrease in adherence to diary use throughout the study, this method might be less suitable for longer trials.

(JMIR Mhealth Uhealth 2020;8(11):e19696) doi:10.2196/19696

KEYWORDS

irritable bowel syndrome; digital diary; smartphone application; mobile phone application; mhealth; e-health; compliance; electronic case report file; patient reported outcome measures; peppermint oil; PERSUADE study.

Introduction

Irritable bowel syndrome (IBS) is a highly prevalent chronic disorder of brain-gut interaction characterized by recurrent abdominal pain and altered bowel habits [1]. Since well-defined organic causes and validated biomarkers for IBS are lacking, patient-reported outcome measures are crucial in assessing treatment response. Accordingly, drug regulatory authorities currently recommend using end-of-day symptom scores in IBS trials to measure drug efficacy [2,3]. Diaries are generally considered to be suitable to measure end-of-day gastrointestinal symptom scores and have the ability to capture symptom variability over time [4]. The validity and reliability of paper diaries, however, may be impeded by fake adherence [5] (ie, falsifying or backfilling written answers outside of the proposed time window [6]). The gap between reported and actual adherence to paper diaries has been shown to be as large as 80% in some studies [7]. Because backfilling introduces considerable recall and ecological bias [8], using paper diaries can distort trial results, which can ultimately lead to incorrect conclusions about treatments. Efficacy endpoints in clinical trials should, therefore, preferably not be assessed by paper diaries.

Recent technological advancements and the widespread availability of smartphones have given rise to numerous health-related apps and electronic diaries in the last decade [9-12], both in clinical and research settings. Digitalized data collection provides several advantages over a paper-based data collection as it results in higher data entry quality and more efficient data handling [13]. For example, responses can be verified automatically by built-in response requirements, routing, and data validation, and manual data transcription can be omitted. More importantly, data entry for previous days can be prevented, and all entries can be given a date- and time-stamp, generating more valid (momentary) results and allowing assessment of actual adherence to the diary. Studies that have implemented electronic diaries have reported excellent adherence, ranging from 76%-100% [5,14,15].

These advantages encouraged our group to implement a digital data-collection framework and develop a smartphone app that can be used as a digital symptom diary. This diary was used to collect Food and Drug Administration (FDA) recommended efficacy outcomes in our randomized placebo-controlled clinical trial on the efficacy of peppermint oil in IBS called the PERSUADE study [16]. This observational study describes the development and evaluates the performance of the overall digital framework used for data collection in that clinical trial. Within the realm of IBS trials, the use of a digital symptom diaries is relatively new; most previous studies have not reported adherence for the assessment method used, and data on adherence in other populations cannot necessarily be extrapolated to IBS. Therefore, our primary aim was to evaluate the performance of a custom-made digital symptom diary in patients with IBS, in particular by assessing patients' adherence. Since patient characteristics can impact adherence [17,18], our secondary aim was to identify sociodemographic and clinical patient characteristics associated with adherence.

Methods

Overview

The study was based on data from the PERSUADE study [16]. This was a randomized double-blind placebo-controlled trial (clinicaltrials.gov; NCT02716285) conducted in 4 hospitals located throughout the Netherlands (Multimedia Appendix 1; Figure S1). The study protocol was approved by the Maastricht University Medical Center+ Ethics Committee. All study procedures were performed in compliance with Good Clinical Practice Guidelines and according to the revised Declaration of Helsinki [19]. All participants gave written informed consent prior to participation.

The study design of the PERSUADE has been described in detail elsewhere [16]. In brief, the primary aim was to investigate the efficacy of peppermint oil—a conventional small-intestinal release formulation and a novel ileocolonic release formulation—in patients with IBS. To this end, patients between 18-75 years of age, who fulfilled the Rome IV criteria for IBS and had a mean worst abdominal pain score of at least 3 on an 11-point rating scale (0, no pain; 10, worst possible pain) during a 14-day pretreatment period were included. Participants were randomized to placebo, small-intestinal release peppermint oil, or ileocolonic release peppermint oil for an 8-week treatment period.

Data were collected using a customized framework for digital data collection, specifically designed and developed for the trial, consisting of (1) a digital symptom diary (smartphone app); (2) an electronic case report file (eCRF, Castor EDC); (3) web-based patient questionnaires (Castor EDC); and (4) a planning tool (Ldot). During the 14-day pretreatment and the 8-week treatment period, patients were instructed to register symptoms daily in the digital symptom diary. Study visits and telephone follow-up telephone interviews were documented in the eCRF. Patients were asked to complete several web-based questionnaires at different time-points within the study duration. The complete list of inclusion criteria and study overview with timing of the questionnaires is given in the Multimedia Appendix 1. Primary efficacy results of the PERSUADE study have been described elsewhere [16].

Digital Symptom Diary: Smartphone App

For the digital symptom diary, an electronic smartphone app was developed by the Center for Data and Information Management at Maastricht University (MEMIC), in close collaboration with the investigators. The app was programmed using Xamarin, a framework to develop cross-platform apps using C sharp programming. The PERSUADE app supports Android and iOS devices. A Maastricht University industrial designer designed the visual content. A MEMIC team of data managers and researchers of the Maastricht University Medical

Center+ Neurogastroenterology group tested the app and provided feedback throughout several phases of development. Additionally, a patient was asked to use the diary and provide feedback regarding its user friendliness. Patient inclusion commenced once a version was reached that all agreed on.

The app's home screen consisted of 3 main elements: the daily end-of-day symptom questionnaire, a medication list, and the Bristol stool chart questionnaire (Figure 1). The end-of-day symptom questionnaire included one main question to assess the primary outcome (in accordance with FDA guidelines): "How would you rate your abdominal pain today? Think about the worst abdominal pain today" (11-point numerical rating scale) (Figure 2). The daily symptom questionnaire was accessible between 6 PM and 12 PM and was unavailable outside this time window, to avoid premature completion. Other daily questions were related to "need of rescue medication" and "adverse events experienced." If a patient had not completed the daily entry before 10 PM of that particular day, one push notification was sent. At the end of each week, the end-of-day questionnaire consisted of additional questions regarding abdominal discomfort, abdominal bloating, abdominal cramping, belching, nausea, and urgency during the last week (11-point scale numerical rating scale). It was not possible to enter data for previous days, and participants could not review prior entries. Automated routing, response requirements, and real-time data verification were built in to increase data quality and completeness.

The medication list was used once to register all regular medications. Patients were asked to keep their concomitant medication use as stable as possible. However, if alterations were needed, they were able to delete, add, or change the dosage of nongastrointestinal drugs.

The Bristol stool form scale was used to register all bowel movements (Figure 3). There was no minimum or maximum number of registrations per day.

Figure 1. Home screen of the smartphone app.

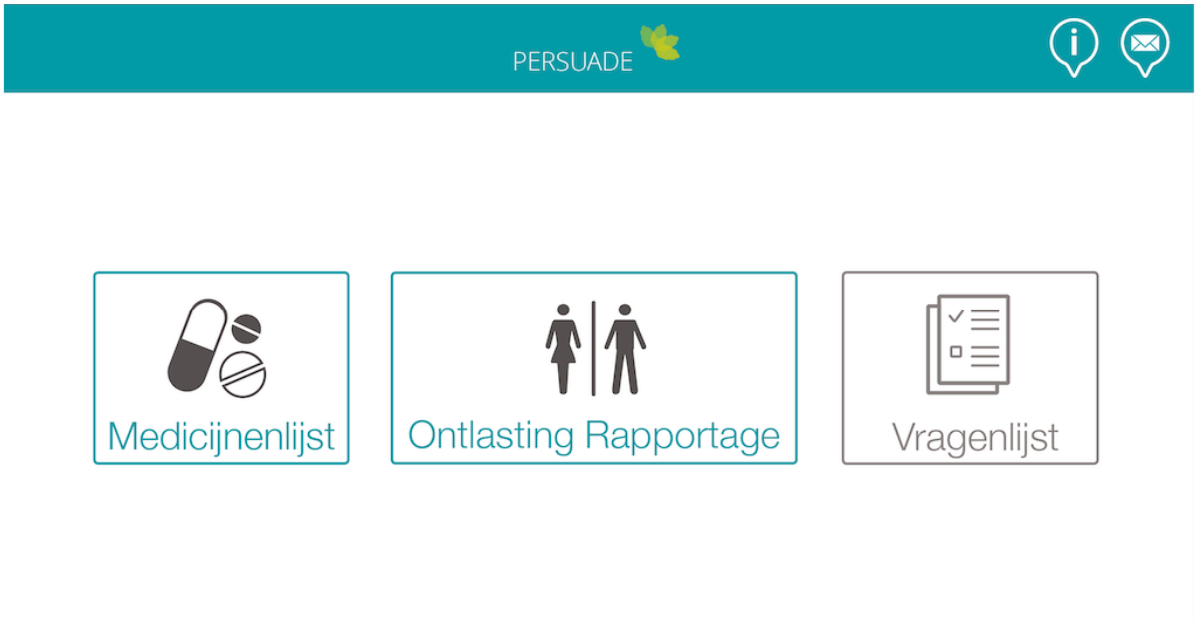


Figure 2. Primary outcome measure.

Figure 3. Bristol Stool Form Scale.

All patients received extensive verbal and written instructions during the screening visits on how to use the app and were encouraged to contact the researchers if the app crashed or otherwise did not function properly. A personalized username and password were provided for access to the app. Patients were instructed to enable automatic updates of the app to ensure the most recent version was used. If a patient did not own a smartphone or tablet, a device was provided. During the complete pretreatment and treatment periods, an alert system in the planning tool notified the investigators when a patient had failed to submit 3 or more daily entries. In addition, the development team received automated notifications of app crashes.

Web-Based Patient Questionnaires

At randomization, at 2, 4, 6, and 8 weeks of treatment, and at 3 and 6 months of follow-up after treatment ended, patients were asked to fill out web-based questionnaires. We chose not to implement these into the digital diary because of the large number of questions. Included were a questionnaire regarding demographics and lifestyle and validated questionnaires regarding symptom severity (IBS Symptom Severity System), quality of life (IBS Quality of Life, EuroQoL EQ-5D-5L), comorbid symptoms of anxiety and depression (General Anxiety Disorder 7, Patient Health Questionnaire 9), and health care utilization and productivity loss (Medical Consumption Questionnaire, Productivity Cost Questionnaire). Patients received invitations via email containing an HTML-link to the electronic environment. If a patient had not completed the

questionnaire within 2 days, 2 automatic reminders were sent via email. Automated routing, response requirements, and real-time data verification were built in to increase data-quality and completeness.

eCRF

During the study visits and telephone follow-up calls, investigators documented all findings in a cloud-based eCRF. The eCRF forms were built by the first author with input from the other authors and contained items regarding demographics, Rome IV diagnostic criteria for IBS, history and physical examination, adverse events, and general wellbeing. Investigators were given unique usernames and passwords to view and add data for their respective inclusion centers. To achieve registration uniformity, the investigators were trained on how to enter data, and additional step-by-step instructions were given in a standard operating procedure document. Real-time automated data verification and corresponding pop-up notifications were built in to prevent typing errors or other erroneous entries. Automated routing of questions and response requirements ensured that correct items were displayed and filled in. An audit trail enabled tracking of all data changes.

Ldot Planning Tool

Ldot is a web-based tool developed by the Center for Data and Information Management at Maastricht University and was used to monitor study logistics. All personal patient data were entered into Ldot, and the app supported the study workflow by indicating when each study event (eg, randomization, follow-up call, etc) needed to take place for each patient. Ldot was able to communicate with the digital diary and the web-based questionnaires. For example, all email invitations for the questionnaires were sent automatically via Ldot. Patients' adherence to the diary and web-based questionnaires could be monitored within Ldot and investigators were notified if patients failed to complete 3 consecutive days in the diary. Investigators were also notified if patients failed to complete a web-based questionnaire after a reminder was given. To guarantee the anonymity and quality of research data, no research data could be entered into Ldot. Investigators could view and add personal data for their respective inclusion centers. There was no possibility of viewing data from other inclusion centers, except for the coordinating investigator (first author) who had access to all data. An audit layer of the app tracked and stored information of all changes.

Storage, Servers, and Privacy

All software and data storage complied with international ISO27001, ISO9001, good clinical practice guidelines, and Dutch NEN7510 guidelines. Electronic diary data, web-based-questionnaire data, eCRF data, and sensitive personal data (Ldot) were all stored on different (nonconnected) servers. Several back-ups were made per day. Access to the servers was and will be restricted, with 24-hour on-site surveillance. Data will be stored for 15 years after study completion.

Outcome Measures

The primary outcome of the current study was patients' adherence to the digital symptom diary, defined as the mean percentage of entries and calculated by dividing the number of completed entries by the number of minimal requested entries (total number of days in study). Patients were instructed to complete a diary entry on all consecutive days during the 14-day pretreatment and 56-day treatment period, or all days until discontinuation with the study.

Secondary outcomes were change in mean adherence per week over time, sociodemographic characteristics, clinical patient characteristics associated with adherence, time of diary completion, and difference in adherence between patients who were defined as responders to treatment versus nonresponders. Potential data loss and critical evaluation points were considered to explore the overall feasibility of a smartphone app as a primary data-collection tool in a randomized controlled trial. Other secondary outcomes were patients' adherence to and completeness of the additional web-based questionnaires and investigators' adherence to and completeness of the eCRF.

Statistical Analysis

Statistical analyses were carried out using SPSS statistical software (version 25.0 for Macintosh; IBM Corp). Data are expressed as mean and standard deviation or as number and percentage. Multivariable linear regression analysis was used to investigate the association between baseline patient characteristics and adherence to the digital diary, adjusting for minimization variables (age, gender, IBS subtype, inclusion center, and treatment group). A repeated measures analysis of variance was performed to assess the influence of time (weeks) on adherence. If the Mauchly test indicated that the sphericity assumption was not met, Greenhouse-Geisser corrected results were reported. A $P<.05$ (2-sided) was considered statistically significant.

Results

General

Overall, 190 patients were randomized. One patient was randomized erroneously (ie, without fulfilling all inclusion criteria). Therefore, 189 patients (age: mean 34.0, SD 13.3 years; female: 147/189, 77.8%; male: 42/189, 22.2%) were analyzed (n=64 in the placebo group, n=62 in the small intestinal release peppermint group, n=63 in the ileocolonic release peppermint oil group). Of the 189 patients, 95.8% (181) were Caucasian and 4.2% (8) were of mixed descent. Most patients (109/189, 57.7%) were recruited from a primary care setting. Eleven patients withdrew from the study during the treatment period (data until discontinuation were included in the analyses). Baseline characteristics are presented in Table 1. During recruitment, only a single patient stated the digital data collection as a reason not to participate.

Table 1. Summary of patient demographic and baseline characteristics.

Characteristic (N=189)	Value
Age (years)	
Mean (SD)	34.0 (13.3)
Range	18-70
Gender, n (%)	
Male	42 (22.2)
Female	147 (77.8)
Education level, n (%)	
No education	1 (0.5)
Low	15 (7.9)
Moderate	80 (42.3)
High	93 (49.2)
Setting, n (%)	
Primary care	109 (57.7)
Secondary care	41 (21.7)
Combined secondary & tertiary care	39 (20.6)
IBS[a]-subtype[b], n (%)	
Diarrhea	83 (43.9)
Constipation	42 (22.2)
Mixed	40 (21.2)
Undefined	24 (12.7)
IBS severity[c]	
Score, mean (SD)	276.5 (71.9)
Mild, n (%)	15 (7.9)
Moderate, n (%)	100 (52.9)
Severe, n (%)	74 (39.2)
IBS Quality of Life score[d], mean (SD)	73.0 (15.1)
EQ-5D-5L utility score[e], mean (SD)	0.7 (0.2)
Psychological comorbidities[f], mean (SD)	
Anxiety	5.4 (4.3)
Depression	6.8 (4.5)

[a]IBS: irritable bowel syndrome.

[b]Determined in a face-to-face interview (according to Rome IV criteria).

[c]The Irritable Bowel Syndrome (IBS) Symptom Severity System consists of 5 items with a maximum score of 100, higher scores indicate more severe symptoms.

[d]The IBS Quality of Life questionnaire consists of 34 items with a 5-point Likert scale (1=good, 5=worse).

[e]The EuroQol-5D-5L measures 5 dimensions of quality of life. Raw scores are transformed to utility scores [20], which vary from 1 (perfect health) to 0 (death).

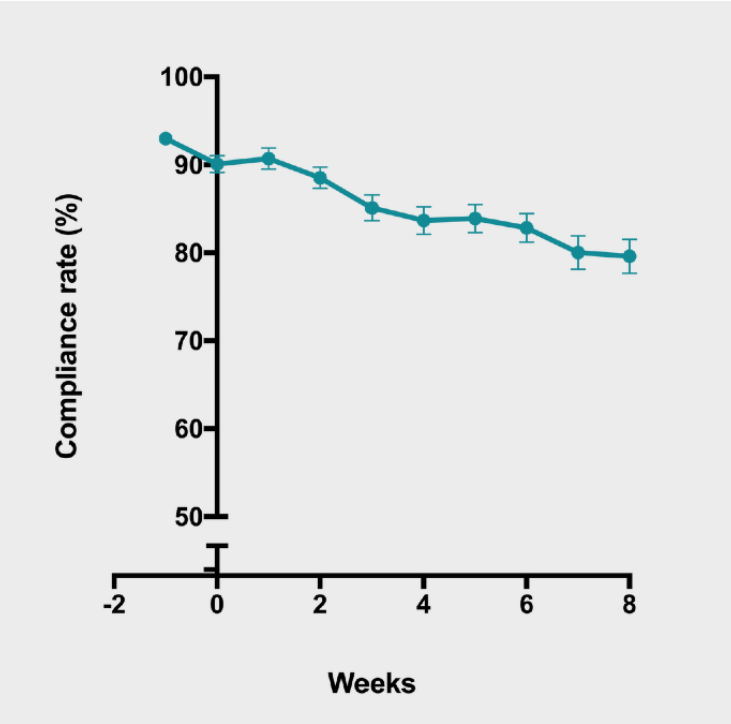
[f]The General Anxiety Disorder–7 consists of 7 items with a 4-point response scale (0=not at all, 3=almost every day). The Patient Health Questionnaire–9 consists of 9 items with a 4-point response scale (0=not at all, 3=almost every day).

Patients' Adherence to the Digital Symptom Diary

Most patients used their own smartphones, but 4 out of 189 patients needed a device provided by the investigators. Patient adherence to the daily digital symptom diary was excellent during the entire study period, reflected by a mean completion rate of 87.9% (SD 9.4%), 91.5% (SD 9.2%), and 86.9% (SD 10.8%), during all 70 days of study duration, the 14-day pretreatment period, and the 8-week treatment period, respectively. Adherence during the treatment period did not

differ significantly for treatment groups compared to that of the placebo (placebo: mean 87.2%; small-intestinal release: 88.3%; *P*=.67), and for the ileocolonic release peppermint oil (mean 87.2%; *P*=.33). Adherence did not differ between patients that were clinical responders to treatment (mean 88.0%) versus patients who were nonresponders (mean 86.2%). Over the complete study period of 70 days, a significant decrease in mean weekly patient adherence to the end-of-day questionnaire was found ($F_{5.9,\ 1114.9}$=15.5, *P*<.001) (Figure 4). Nevertheless, adherence was still good at the end of the study (mean 79.6%, SD 26.6%) (Figure 4).

Figure 4. Adherence.



When exploring independent baseline predictors for adherence, the combined regression model that included all minimization variables (age, gender, IBS subtype, inclusion center), treatment group, baseline IBS symptom severity, anxiety and depression scores, and education level showed that 1 of the 4 inclusion centers (center C; Multimedia Appendix 1, Figure S2), (regression coefficient B –10.04, 95% CI –19.51 to –0.56, *P*=.04) and anxiety scores at baseline (B –0.59, 95% CI –1.12 to –0.06, *P*=.03) were negatively associated with adherence throughout the study. Indeed, when comparing adherence between the different inclusion centers, adherence in center C was lowest (mean 82.3%, SD 12.5%), compared with adherence in centers A (mean 88.3%, SD 9.2%), B (mean 84.7, SD 14.0%), and D (mean 91.4%, SD 7.7%). There was no statistically significant effect of gender (*P*=.84), age (*P*=.22), or education level (lower education level: *P*=.58, middle education level: *P*=.46, versus high education level) on adherence. Mean time of completing the end-of-day symptom diary was 9:46 PM (ie, 14 minutes before receiving the push notification).

Feasibility of a Smartphone App as Primary Data-Collection Method

Several technical issues were noted by the investigators or reported by patients. In most cases, the cause was found, and the issue was resolved by the app development team without data loss. Encountered hurdles included difficulties installing the app during the screening visit due to connectivity failure, not receiving reminder notifications, inaccurate visual scaling of questionnaires on smaller smartphone screens, updates of Android or iOS operating systems that interfered with prior processes, and connectivity failure due to server maintenance. All documented technical issues and their short-term consequences are presented in Table 2.

Table 2. Technical difficulties and consequences with regard to the digital symptom diary.

Description of technical issue	Patients affected, n	Consequence and solution, if applicable
Low internet connectivity hindered installation of the app during the screening visit	15	In most cases, the problem was solved by moving to a location with better internet connectivity or by postponing the installation to a later time.
Not receiving push-notifications as a reminder to complete the end-of-day questionnaires	12	In many cases, patients would complete the questionnaire regardless of receiving the notification. However, the exact effect is unknown, and it may have negatively impacted adherence during days that no notification was received. In most cases, the problem could be resolved by changing the telephone settings (eg by ignoring battery optimizations). In 2 cases in which the issue could not be resolved, reminders were given during the study period by setting the alarm of the device at 10 PM. In the short period during which it was unknown how many devices were affected, additional text messages were sent as reminders.
Incomplete views of the questions due to a too large scaling on smaller smartphone screens.	8	The issue was resolved by adjusting the scaling in the app during updates. Because only a few letters were not depicted correctly and because all participants had received a manual that included the actual questions asked, the negative effect of short-term scaling issues was estimated to be negligible.
iOS or Android updates that interfered with prior settings of the app	0	The issue did not lead to missing data because the small bugs did not shut down the app. The development team provided updates that resolved the issues as soon as possible.
Maintenance of the hosting server	21	The issue led to missing data of one complete day (ie, the day on which the maintenance took place) in all but 2 patients who were included at the time of the maintenance.

Web-Based Questionnaires: Patients' Adherence and Completeness

Adherence to the web-based questionnaires was also excellent. One patient did not complete the questionnaires at the end of the treatment period; all others completed all questionnaires until the end of the study or until discontinuation (n=11 discontinued the study). Halfway through the study duration, however, a routing error in one questionnaire became apparent. Although this mistake was corrected immediately, the error had already led to missing data for that particular question in 23.3%-54.0% of all patients, depending on measurement moment (Table S1 in Multimedia Appendix 1). No missing items were found in other questionnaire items.

Investigators' Adherence to the eCRF

Adherence of the investigators to the eCRF was excellent with a completion rate of more than 99%. In total, there were 27 patients with at least 1 missing variable in the case report file, 11 of whom discontinued the study during the treatment period (the missing values comprehend follow-up calls that were not conducted). The remaining 17 cases with missing data were because of missed follow-up calls (in 11 cases, 1 follow-up (out of 3) was missed), not registering if additional information about the 6-month follow-up period was given, not registering the date of the last menstruation, not registering if the general practitioner was informed about participation in the study, or not registering the number of capsules that were reported not to be taken during one of the follow-up calls.

Discussion

The results of this study demonstrate that patients' adherence to the end-of-day questionnaire in the digital symptom diary was excellent, with a mean completion rate of 87.9% over 70 days of study duration. The total proportion of missing data and data loss due to technical issues of the app was small, indicating that it is safe and realistic to use the app as a primary data-collection method. Furthermore, patients' adherence to the web-based questionnaires and investigators' adherence to the eCRF were also outstanding with completion rates of more than 99%.

In terms of electronic diary usage in clinical trials, the adherence rate found in this IBS study is at least comparable to or higher than previously reported rates [5,14,15]. Most people (90.3%) in the Netherlands own a mobile phone [21]. Only a few patients (n=4) needed a device from the investigator team to participate in the study, and only a single patient stated digital data collection as a reason not to participate. Mean adherence to the digital symptom diary decreased by 11% from the first week of the pretreatment period to the last week of the 8-week treatment period (Figure 4). A slight decrease in adherence to the diary during a study period (ie, logging fatigability) is not uncommon and has also been observed in other studies investigating digital diaries [5,22]. Regarding the usage of digital diaries in randomized controlled trials to assess treatment response (according to FDA-recommended definitions) in IBS patients specifically, we are aware of one recent IBS study [23] that applied an electronic diary to assess treatment effect. However, a direct comparison with this study was not possible, as details on the type of device, app, or adherence to the diary were not provided.

With regard to sociodemographic and clinical patient characteristics associated with completing the daily entries in the diary, we found no evidence of a statistically significant effect of gender (P=.84), age (P=.22), or education level (lower education level: P=.58, middle education level: P=.46, versus high education level) on adherence. This differs from the results of some prior studies and meta-analysis [15,17] that observed, for example, a statistically significant positive effect of age on adherence. Our interpretation is that this may be caused by the

relatively young patient population in the current study. We observed that patients with higher anxiety scores had lower adherence to the digital symptom diary. These data are in line with those of Aaron et al [24], showing that participants with higher stress levels may have lower completion rates. Interestingly, a negative association was found between one inclusion center and adherence. All 4 inclusion centers were located in urban areas but with a wide geographical spread throughout the Netherlands as shown in Multimedia Appendix 1 (Figure S2). The center with the negative association (center C) was in the most urban and populated area (ie, the Amsterdam-The Hague-Rotterdam-Utrecht urban agglomeration). No obvious demographic or baseline differences were observed between study populations in different inclusion centers. No association was found between the lower adherence and the investigator by whom the instructions were given. Although the reason for lower adherence of patients included in this center is unclear, religious and cultural backgrounds of inhabitants of this agglomeration may have differed from those of the inhabitants of other geographical areas [25-27]. Nevertheless, overall adherence during the treatment period in this inclusion center was still good (mean 82.3%, SD 12.5%).

In terms of technical issues arising during the study, minor bugs occurring as a consequence of ever evolving smartphones and operating systems are practically inevitable. It is our experience, therefore, that continuous maintenance and software updating by a development team is crucial to avoid data loss and potential agitation of the study participant due to app malfunctioning. Consequently, the feasibility of using a smartphone app as a primary data-collection method depends to a large extent on skills and availability of development team staff, and research groups should check if appropriate support is available before opting for such methods.

Many high-quality IBS trials have used interactive voice response systems as the primary data collection method [28-31]. In spite of this frequent use, the interactive voice response systems used in IBS trials have not been described in detail, thereby hampering replication and implementation of the methodology in other trials. For comparison with our methodology, we therefore depended on what is known about interactive voice response systems in general. Akin to a digital symptom diary, the interactive voice response systems method allows control of time-windows in which surveys should be completed, provides automated time-stamps to answers, performs data verification and validation, follows a predefined routing schema, enables automatic reminders, collects and stores data in real time, and leads to an overall consistent survey administration. In addition, both methods equally depend on telephone- or internet-service and require staff to program and maintain the software. A potential advantage of interactive voice response systems over those of digital diaries is that it does not depend on literacy skills of the participant. An interactive voice response systems may also need fewer software updates than what is required by smartphone apps due to the high pace of smartphone operating system updates. Potential disadvantages of the interactive voice response systems compared with digital symptom diaries are (1) the inability to get clarification during the survey, whereas a digital symptom diary can have built-in optional clarification of questions; (2) not all interactive voice response systems are equipped with speech recognition; open-ended questions then require transcription by a data manager; (3) the quality of open-ended question recordings depends on enunciation, background noise, and connection; and (4) usage of the interactive voice response systems requires extensive participant training and could be less user friendly [32]. As for patient adherence to the interactive voice response systems, this was reported by only one recent IBS trial [30]; they reported a mean adherence rate of 71% and 73% in the 2 groups examined, when adherence was defined as completing at least 80% of the scheduled calls to the interactive voice response systems. Adherence to the interactive voice response systems in that study [30] was thus notably lower than adherence to the digital symptom diary found in this study.

This study described the overall framework for digitalized data collection used in the PERSUADE study. In addition to the digital symptom diary, the electronic framework used in this drug trial consisted of web-based patient questionnaires and an electronic CRF to collect additional secondary outcomes. A troublesome issue that occurred was a routing error in one of the questionnaires that was discovered too late and had already led to a high proportion of missing data (Table S1 in Multimedia Appendix 1). This applied to only a single question, but routing errors can have potentially disastrous consequences. As such, investigators and data managers should take appropriate care and time when testing questionnaires. Data exports should, furthermore, be examined in an early testing phase and preferably by more than one investigator and data manager. Similar to the diary, the web-based questionnaires and the eCRF featured built-in routing of questions, data validation, and response requirements to encourage data quality and completeness. Overall, these steps allowed for guaranteed standardized data collection with completeness of more than 99% for the web-based questionnaire and eCRF items.

Additional advantages of the combined framework for digitalized data-collection are (1) the ability to monitor patients and their adherence; (2) a reduction in paperwork and physical archiving (eg, in this study the paperwork was reduced to one single informed consent file); (3) manual data transcription can be omitted as research data enter the database immediately; (4) the possibility to adjust and individualize the smartphone app, eCRF, and web-based questionnaires according to the needs of each particular study, and (5) more accurate and standardized data reporting since no error-prone re-entry is necessary. The described framework for digital data collection can, therefore, be employed in different studies investigating different disease entities.

Our findings should be interpreted in light of some potential limitations. First, the study was not primarily designed for the analysis of adherence to the digital symptom diary but for measuring the main clinical outcome [16]. However, since almost 200 patients were included, the sample size was sufficiently large to estimate adherence with enough precision. Second, adherence rate was not assessed within a controlled trial with a more traditional method of data collection (ie, paper-and-pencil diaries, interactive voice response systems) as a comparison. However, the rapid diffusion toward digital

approaches in health care and clinical research renders such comparisons less meaningful from a practical point of view as the use of digital techniques become inevitably ubiquitous. In addition, it is unlikely that these traditional approaches to data collection would result in higher adherence than those observed here. Another limitation was that patient satisfaction with the digital diary or web-based questionnaires was not quantified by means of a questionnaire. In this study, the feasibility of the used framework was evaluated primarily on the basis of patients' and investigators' adherence and the proportion of complete data, whereas quantified patient satisfaction was not taken into account. However, a low patient satisfaction would have likely led to a lower adherence, and thereby, a higher proportion of missing data. Therefore, although helpful, it is unlikely that applying such a questionnaire would have altered our main findings.

In this IBS drug trial, the use of a smartphone app as a digital symptom diary to assess treatment response was found to be highly feasible and resulted in high quality data collection with excellent patient adherence of more than 86% during the complete study period. The combination of the digital diary with the eCRF, planning tool, and web-based questionnaires led to overall standardized state-of-the art data collection with excellent completeness and can be used as a framework for future randomized controlled trials. Due to the slight decrease in patient adherence to the digital diary throughout the study, caution is needed when using such methods in long-term studies. Although this framework was designed for IBS clinical trials, the results reported here are of added value to a far broader range of disorders for which the collection of patient-reported outcome measures is required. Future studies should preferably include a control group, for example, a group using the interactive voice response system or a group using the app without receiving reminder notifications, to compare adherence and to ascertain specific factors driving high adherence.

Acknowledgments

We thank all patients who participated in the PERSUADE study; staff members of the PERSUADE team in the 4 participating centers (Maastricht UMC: A Bours, N Aendekerk, D Janssen, G Van Hooff; Hospital Gelderse Vallei: B Witteman, G Homans, A Van de Vendel; Alrijne hospital Leiden: C Clemens, M Oosterveer; Medical Center Leeuwarden: A Westendorp-Ijdema, A de Ruiter-van der Ploeg, G Tack-Blokker, H Lebbink); other members of the PERSUADE study team (J Brouwers, J Tack, H Frijlink, N de Wit, J Muris, B Essers); staff members of Center for Data and Information Management at Maastricht University (D Veldman, J Pisters, L Linden, D de Jong); the Dutch IBS patient federation; and Will Pharma SA employees (K Bouwens, AH Verhage, F Bunel, K Abildayeva).

Authors' Contributions

ZZRMW was responsible for study concept and design, development of online questionnaires, data collection, data cleaning, data analysis, data interpretation, and manuscript writing. KGEH was responsible for development and maintenance of the smartphone app and planning tool, data cleaning. AAMM was responsible for obtaining funding, study concept and design, data interpretation, and constructive review of manuscript. ABAQ was responsible for data collection and data analysis. BW was responsible for statistical support and constructive review of manuscript. LV was responsible for development of online questionnaires, data-collection, and constructive review of the manuscript. PELMR was responsible for development and maintenance of the planning tool and online questionnaires, as well as data cleaning. DMAEJ was responsible for study concept and design, data interpretation, and constructive review of manuscript. DK was responsible for obtaining funding, study concept and design, data interpretation, and constructive review of manuscript. All authors approved the final manuscript.

Conflicts of Interest

AAMM and DK have received a ZonMw, The Netherlands Organization for Health Research and Development (Dutch government) health care efficiency grant for the execution of the PERSUADE study (grant number 836031017). AAMM and DK have received an unrestricted research grant from Will Pharma SA, which also supported ZZRMW and LV in attending a scientific meeting. The study design, data collection, analysis, and interpretation were performed by the academic authors without industry involvement. The employer of DK and AAMM have an agreement with Will Pharma SA regarding the exploitation of a potential market authorization of the ileocolonic formulation of peppermint oil for IBS. AAMM and DK have received research funding from Allergan and Grünenthal involving a different smartphone app for IBS. KGEH, ABAQ, BW, and PELMR have no conflicts of interest. Part of the work of DJ is financed by Grant Top Knowledge Institute (Well on Wheat), the Carbokinetics program as part of the NOWO-CCC Partnership program, and H2020 848228/DISCOvERIE.

Multimedia Appendix 1
Supplementary material.
[DOCX File , 221 KB - mhealth_v8i11e19696_app1.docx]

References

1. Lovell RM, Ford AC. Global prevalence of and risk factors for irritable bowel syndrome: a meta-analysis. Clin Gastroenterol Hepatol 2012 Jul;10(7):712-721.e4. [doi: 10.1016/j.cgh.2012.02.029] [Medline: 22426087]
2. Guideline on the evaluation of medicinal products for the treatment of irritable bowel syndrome. European Medicines Agency. 2013. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-medicinal-products-treatment-irritable-bowel-syndrome-revision-1_en.pdf [accessed 2020-10-15]
3. Guidance for Industry - Clinical Evaluation of Drugs for Treatment. Center for Drug Evaluation Research. 2012. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/irritable-bowel-syndrome-clinical-evaluation-products-treatment [accessed 2020-10-15]
4. Stone AA, Schwartz JE, Broderick JE, Shiffman SS. Variability of momentary pain predicts recall of weekly pain: a consequence of the peak (or salience) memory heuristic. Pers Soc Psychol Bull 2005 Oct;31(10):1340-1346. [doi: 10.1177/0146167205275615] [Medline: 16143666]
5. 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]
6. Lauritsen K, Degl' Innocenti A, Hendel L, Praest J, Lytje MF, Clemmensen-Rotne K, et al. Symptom recording in a randomised clinical trial: paper diaries vs. electronic or telephone data capture. Control Clin Trials 2004 Dec;25(6):585-597. [doi: 10.1016/j.cct.2004.09.001] [Medline: 15588745]
7. 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]
8. Mujagic Z, Keszthelyi D, Aziz Q, Reinisch W, Quetglas EG, De Leonardis F, et al. Systematic review: instruments to assess abdominal pain in irritable bowel syndrome. Aliment Pharmacol Ther 2015 Nov;42(9):1064-1081 [FREE Full text] [doi: 10.1111/apt.13378] [Medline: 26290286]
9. Riaz MS, Atreja A. Personalized technologies in chronic gastrointestinal disorders: self-monitoring and remote sensor technologies. Clin Gastroenterol Hepatol 2016 Dec;14(12):1697-1705 [FREE Full text] [doi: 10.1016/j.cgh.2016.05.009] [Medline: 27189911]
10. 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]
11. Toivonen KI, Zernicke K, Carlson LE. Web-based mindfulness interventions for people with physical health conditions: systematic review. J Med Internet Res 2017 Aug 31;19(8):e303 [FREE Full text] [doi: 10.2196/jmir.7487] [Medline: 28860106]
12. Kheirkhahan M, Nair S, Davoudi A, Rashidi P, Wanigatunga AA, Corbett DB, et al. A smartwatch-based framework for real-time and online assessment and mobility monitoring. J Biomed Inform 2019 Jan;89:29-40 [FREE Full text] [doi: 10.1016/j.jbi.2018.11.003] [Medline: 30414474]
13. 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]
14. Burton C, Weller D, Sharpe M. Are electronic diaries useful for symptoms research? A systematic review. J Psychosom Res 2007 May;62(5):553-561. [doi: 10.1016/j.jpsychores.2006.12.022] [Medline: 17467410]
15. Morren M, van Dulmen S, Ouwerkerk J, Bensing J. Compliance with momentary pain measurement using electronic diaries: a systematic review. Eur J Pain 2009 Apr;13(4):354-365. [doi: 10.1016/j.ejpain.2008.05.010] [Medline: 18603458]
16. Weerts ZZRM, Masclee AAM, Witteman BJM, Clemens CHM, Winkens B, Brouwers JRBJ, et al. Efficacy and safety of peppermint oil in a randomized, double-blind trial of patients with irritable bowel syndrome. Gastroenterology 2020 Jan;158(1):123-136. [doi: 10.1053/j.gastro.2019.08.026] [Medline: 31470006]
17. Ono M, Schneider S, Junghaenel DU, Stone AA. What affects the completion of ecological momentary assessments in chronic pain research? an individual patient data meta-analysis. J Med Internet Res 2019 Feb 05;21(2):e11398 [FREE Full text] [doi: 10.2196/11398] [Medline: 30720437]
18. Ernsting C, Stühmann LM, Dombrowski SU, Voigt-Antons J, Kuhlmey A, Gellert P. Associations of health app use and perceived effectiveness in people with cardiovascular diseases and diabetes: population-based survey. JMIR Mhealth Uhealth 2019 Mar 28;7(3):e12179 [FREE Full text] [doi: 10.2196/12179] [Medline: 30920383]
19. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
20. M Versteegh M, M Vermeulen K, M A A Evers S, de Wit GA, Prenger R, A Stolk E. Dutch tariff for the five-level version of EQ-5D. Value Health 2016 Jun;19(4):343-352 [FREE Full text] [doi: 10.1016/j.jval.2016.01.003] [Medline: 27325326]
21. Internet; toegang, gebruik en faciliteiten. Statline. 2018. URL: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83429NED/table?fromstatweb
22. Zia J, Schroeder J, Munson S, Fogarty J, Nguyen L, Barney P, et al. Feasibility and usability pilot study of a novel irritable bowel syndrome food and gastrointestinal symptom journal smartphone app. Clin Transl Gastroenterol 2016 Mar 03;7:e147 [FREE Full text] [doi: 10.1038/ctg.2016.9] [Medline: 26938478]

23. Lackner JM, Jaccard J, Keefer L, Brenner DM, Firth RS, Gudleski GD, et al. Improvement in gastrointestinal symptoms after cognitive behavior therapy for refractory irritable bowel syndrome. Gastroenterology 2018 Jul;155(1):47-57 [FREE Full text] [doi: 10.1053/j.gastro.2018.03.063] [Medline: 29702118]
24. 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]
25. Arntz R, Wilmots J, De Rooij J. Voor wie Nederland en Vlaanderen wil leren kennen. Diepenbeek: Wetenschappelijk Onderwijs Limburg; 1978.
26. de Jonge H. Ons soort mensen levensstijlen in Nederland. Amsterdam, The Netherlands: SUN; 1997.
27. Wood CW. The History of the Low Countries. History: Reviews of New Books 1999 Jan;28(1):29-30. [doi: 10.1080/03612759.1999.10527769]
28. Lembo AJ, Lacy BE, Zuckerman MJ, Schey R, Dove LS, Andrae DA, et al. Eluxadoline for irritable bowel syndrome with diarrhea. N Engl J Med 2016 Jan 21;374(3):242-253. [doi: 10.1056/NEJMoa1505180] [Medline: 26789872]
29. Pimentel M, Lembo A, Chey WD, Zakko S, Ringel Y, Yu J, TARGET Study Group. Rifaximin therapy for patients with irritable bowel syndrome without constipation. N Engl J Med 2011 Jan 06;364(1):22-32. [doi: 10.1056/NEJMoa1004409] [Medline: 21208106]
30. Rao S, Lembo AJ, Shiff SJ, Lavins BJ, Currie MG, Jia XD, et al. A 12-week, randomized, controlled trial with a 4-week randomized withdrawal period to evaluate the efficacy and safety of linaclotide in irritable bowel syndrome with constipation. Am J Gastroenterol 2012 Nov;107(11):1714-24; quiz p.1725 [FREE Full text] [doi: 10.1038/ajg.2012.255] [Medline: 22986440]
31. Chey WD, Lembo AJ, Rosenbaum DP. Efficacy of tenapanor in treating patients with irritable bowel syndrome with constipation: a 12-week, placebo-controlled phase 3 trial (T3MPO-1). Am J Gastroenterol 2020 Feb;115(2):281-293. [doi: 10.14309/ajg.0000000000000516] [Medline: 31934897]
32. Abu-Hasaballah K, James A, Aseltine RH. Lessons and pitfalls of interactive voice response in medical research. Contemp Clin Trials 2007 Sep;28(5):593-602. [doi: 10.1016/j.cct.2007.02.007] [Medline: 17400520]

Abbreviations

eCRF: electronic case report file
FDA: US Food and Drug Administration
IBS: irritable bowel syndrome
MEMIC: Center for Data and Information Management at Maastricht University

Edited by G Eysenbach; submitted 29.04.20; peer-reviewed by S Schneider, G Iyawa; comments to author 17.07.20; revised version received 30.08.20; accepted 22.09.20; published 02.11.20.

Please cite as:
Weerts ZZRM, Heinen KGE, Masclee AAM, Quanjel ABA, Winkens B, Vork L, Rinkens PELM, Jonkers DMAE, Keszthelyi D
Smart Data Collection for the Assessment of Treatment Effects in Irritable Bowel Syndrome: Observational Study
JMIR Mhealth Uhealth 2020;8(11):e19696
URL: https://mhealth.jmir.org/2020/11/e19696
doi:10.2196/19696
PMID:33030150

Review

Asynchronous mHealth Interventions in Rheumatoid Arthritis: Systematic Scoping Review

Bart F Seppen[1], MD; Pim den Boer[1], BSc; Jimmy Wiegel[1,2], MD; Marieke M ter Wee[2,3], PhD; Marike van der Leeden[1,4], PhD; Ralph de Vries[5], MSc; Martin van der Esch[1,4], Prof Dr; Wouter H Bos[1], MD, PhD

[1]Amsterdam Rheumatology and Immunology Center, Reade, Amsterdam, Netherlands

[2]Department of Rheumatology, VU Medical Center, Amsterdam UMC, Amsterdam, Netherlands

[3]Department of Epidemiology and Biostatistics, Amsterdam Public Health, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, Netherlands

[4]Department of Rehabilitation Medicine, VU Medical Center, Amsterdam UMC, Amsterdam, Netherlands

[5]Medical Library, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Bart F Seppen, MD
Amsterdam Rheumatology and Immunology Center
Reade
Doctor Jan van Breemenstraat 1
Amsterdam, 1056 AB
Netherlands
Phone: 31 202421800
Email: b.seppen@reade.nl

Abstract

Background: Mobile devices such as smartphones and tablets have surged in popularity in recent years, generating numerous possibilities for their use in health care as mobile health (mHealth) tools. One advantage of mHealth is that it can be provided asynchronously, signifying that health care providers and patients are not communicating in real time. The integration of asynchronous mHealth into daily clinical practice might therefore help to make health care more efficient for patients with rheumatoid arthritis (RA). The benefits have been reviewed in various medical conditions, such as diabetes and asthma, with promising results. However, to date, it is unclear what evidence exists for the use of asynchronous mHealth in the field of RA.

Objective: The objective of this study was to map the different asynchronous mHealth interventions tested in clinical trials in patients with RA and to summarize the effects of the interventions.

Methods: A systematic search of Pubmed, Scopus, Cochrane, and PsycINFO was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. Studies were initially screened and later assessed by two independent researchers. Disagreements on inclusion or exclusion of studies were resolved by discussion.

Results: The literature search yielded 1752 abstracts. After deduplication and screening, 10 controlled intervention studies were included. All studies were assessed to be at risk for bias in at least one domain of the Cochrane risk-of-bias tool. In the 10 selected studies, 4 different types of mHealth interventions were used: SMS reminders (to increase medication adherence or physical activity; n=3), web apps (for disease monitoring and/or to provide medical information; n=5), smartphone apps (for disease monitoring; n=1), and pedometers (to increase and track steps; n=1). Measured outcomes varied widely between studies; improvements were seen in terms of medication compliance (SMS reminders), reaching rapid remission (web app), various domains of physical activity (pedometer, SMS reminders, and web apps), patient-physician interaction (web apps), and self-efficacy (smartphone app).

Conclusions: SMS reminders, web apps, smartphone apps, and pedometers have been evaluated in intervention studies in patients with RA. These interventions have been used to monitor patients or to support them in their health behavior. The use of asynchronous mHealth led to desirable outcomes in nearly all studies. However, since all studies were at risk of bias and methods used were very heterogeneous, high-quality research is warranted to corroborate these promising results.

(JMIR Mhealth Uhealth 2020;8(11):e19260) doi:10.2196/19260

KEYWORDS

mobile health; eHealth; digital health; telehealth; telerheumatology; mHealth; web app; smartphone app; activity tracker; rheumatoid arthritis; rheumatology; review; telemonitoring

Introduction

Background

An increase in the prevalence of rheumatoid arthritis (RA), as well as a general shortage of rheumatologists to treat patients with RA, has been described [1,2]. In combination with rising health care costs, this makes structural changes in the Dutch health care system seem inevitable [3]. The challenge set is difficult, as solutions for the increasing health care costs, the rising prevalence of RA, and the subsequent demand on our outpatient clinics will need to follow the treat-to-target guidelines [4]. Integrating mobile health (mHealth) into daily clinical practice may help overcome the challenges the Dutch health care system is facing in the care of patients with RA and other chronic diseases [5-7]. Especially for patients with RA, the anticipated benefits of mHealth use should be evaluated, as adaption of mHealth might be challenging for older patients with RA or patients impeded by hurting joints [8,9].

The World Health Organization defines mHealth as "medical and public health practice supported by mobile devices such as mobile phones, tablets (…)" [7]. Apps for gait analysis, activity tracking, and video consultations and devices for handgrip strength monitoring have been developed and tested in patients with RA [10-13]. Two types of mHealth interventions can be distinguished: synchronous interventions, such as tele- and video consultation (where health care provider and patient are in direct real-time contact), and asynchronous interventions (no direct real-time contact), such as electronic consultations and remote disease activity monitoring through web or smartphone apps [14,15]. Asynchronous mHealth interventions have not received the same degree of attention as synchronous mHealth in RA [16,17], despite the anticipated benefits such as shorter wait times, lower health care usage, and consultations tailored to need [18-21]. So far, it remains unclear what evidence exists for the use of asynchronous mHealth interventions in patients with RA.

Objective

The objective of this scoping review was to map the different asynchronous mHealth interventions tested in clinical trials in patients with RA and to summarize the effects of the interventions. Ultimately, this should help to identify promising implementations and future research opportunities.

Methods

Study Design

We conducted a scoping review of the literature. This type of review is suitable to map the available evidence in new and developing fields. The value of scoping reviews to evidence-based health care and practice lies in the examination of a broader area to identify gaps in the research knowledge base, clarify key concepts, and report on the types of evidence that address and inform practice in the field [22].

Search Methods

A review protocol was developed based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement [23]. Consequently, a comprehensive search of the bibliographic databases PubMed, Embase.com, Ebsco/PsycINFO, Wiley/Cochrane Library, and Scopus was performed by a medical librarian. Databases were searched from inception to November 20, 2019. The following terms were used as index terms or free-text words (including synonyms and closely related words): smartphones, internet, eHealth, mHealth, wearable, apps, rheumatoid arthritis, and tele-rheumatology.

The search was performed without date, language, or publication status restrictions. Duplicate studies were excluded. The full search strategies for all databases can be found in Multimedia Appendix 1.

Inclusion and Exclusion Criteria

The following inclusion criteria were used: (1) the study population comprised only patients with RA, or the majority of the study population consisted of patients with RA, for whom the data were reported separately, (2) the study evaluated an asynchronous mHealth intervention (ie, the health care provider and patient were not in direct synchronous contact), (3) the study type was randomized controlled trial (RCT), randomized controlled crossover trial, or quasi-experimental clinical trial, (4) the study reported outcomes in relation to the mHealth intervention, and (5) the study was published in English as a full-length paper and an original report.

Studies that reported only qualitative outcomes (eg, from focus groups, semistructured interviews, etc) were excluded from the review. References of retrieved studies were screened for additional relevant studies. Interventions that used a web app were taken into account, as these are easily accessible through mobile devices (smartphone and/or tablet) and therefore regarded as mHealth in this review.

Selection of Studies

Initially, the title and abstract were screened independently for eligibility criteria and blinded to each other with the online tool Rayyan [24]. Full-text papers were retrieved for all abstracts that met the inclusion criteria. Disagreements on the inclusion or exclusion of studies were resolved by discussion.

Data Extraction and Categorization

Data were extracted by one reviewer using a standardized template and verified by a second reviewer. The following data were extracted from each included study: the year of the study, the number of participants, patient characteristics, type of study, type of intervention, duration of study and follow-up, outcome measures, univariate outcomes, and statistical significance. Disagreements or discrepancies in data extraction were resolved by discussion.

Quality Assessment of the mHealth Intervention Studies

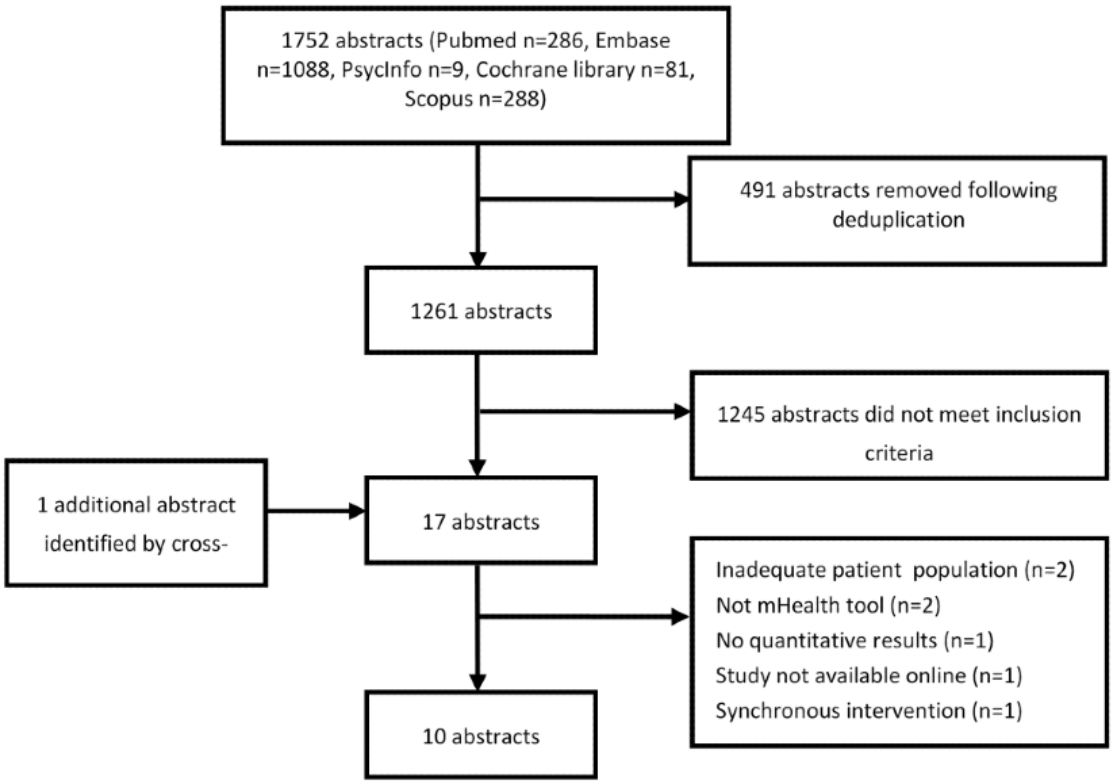
The quality of each study was independently evaluated using the Cochrane Collaboration's tool—a 7-domain tool for assessing risk of bias in which study characteristics are classified as having high, low, or unclear risk of bias [25]. For each study, all items were rated by the same reviewers and substantiated in a Microsoft Excel file (version 2013; Microsoft Corporation). The results were compared and disagreements in results were resolved by discussion and, when necessary, by consultation with a third reviewer.

Results

Study Selection and Inclusion

The search yielded 1752 abstracts (Pubmed: n=286; Embase: n=1088; PsycInfo: n=9; Cochrane library: n=81; and Scopus: n=288). One additional abstract was added after cross-referencing. After deduplication, 1261 abstracts remained, of which 1245 did not meet the inclusion criteria. The main reasons to exclude abstracts were that they were not mHealth studies, the population was not patients with RA, they were conference abstracts, or they were the wrong study type (ie, review or not an intervention study). A total of 17 full-text studies were reviewed and ultimately 10 studies were included (Figure 1). All studies were RCTs, with the exception of a study by Mollard and Michaud [8], which was a nonrandomized controlled study.

Figure 1. Study selection. mHealth: mobile health.



Study Results

A description of the quantitative evidence for the mHealth interventions reported in the included studies is presented in Table 1. In the 10 intervention studies, four different types of interventions were used: SMS reminders (for adherence to medication and physical activity plans) [26-28], web apps (for information and disease monitoring) [13,29-32], smartphone apps (for disease monitoring) [8], and pedometers (for activity tracking) [33]. Measured outcomes and study methods varied among the trials, depending on the aim of the mHealth intervention. All of the studies included between-group comparisons.

Table 1. Evidence for effectiveness of mobile health trials in patients with RA.

Ref[a]	Pts[b] with RA[c], n	Study duration (+ follow-up)	Intervention	Outcome measure	Results[d]		
					Intervention, mean (SD)	Control, mean (SD)	*P* value
[8]	63	6 mths[e]	LiveWith Arthritis smartphone app with optical imaging	Self-efficacy: P-SEMS[f]; PAM[g]	P-SEMS: 2.80[h]; PAM: 6.37[h]	P-SEMS: 1.66[h]; PAM: 2.30[h]	P-SEMS: .04; PAM: .46
[13]	160	12 mths	Physical activity intervention (web app)	Participants that meet Dutch public health physical activity recommendation (%)	6 mths[i]: 38; 12 mths[i]: 26	6 mths[i]: 22; 12 mths[i]: 15	.04; .12
[26]	96	6 mths	SMS reminders	CQR-9[j]	3.32 (5.66)	–0.14 (7.56)	.02
[27]	20	16 wks[k]	Motivational counselling and SMS reminders	Daily sitting time (hours/day)	–0.30 (1.90)	0.15 (1.43)	—[l]
[28]	150	16 wks	Motivational counselling and SMS reminders	Daily sitting time (hours/day)	–1.61 (CI –1.97 to –1.25)	0.59 (CI 0.24 to 0.95)	<.001
[29]	157	2 mths + 2 mths	Access to different sections of a web app: (1) social support, (2) gaming, (3) information, (4) 1 and 2, and (5) control group	Physical activity (β)[m], (minutes); health care use (β), (number of visits); medication overuse (β), (POMI)[n,o]	Physical activity, group 4: β=3.39[p]; health care use, group 2: β=–0.41, group 4[p]: β=–0.33	N/A[q]	Physical activity, group 4: .02; health care use, group 2: .01, group 4: .02
[30]	320	12 mths	Sanoïa (web app)	Patient-physician interaction (PEPPI-5)[r]	0.6 (5.52)	–0.91 (6.08)	.01
[31]	44	12 mths	Telemonitoring with RETE-MARCHE (web app)	Patients with CDAI[s] remission and comprehensive disease control after 1 year (%)	38.1	25	<.01
[32]	108	10 wks + 9 mths	Educational modules for improving self-efficacy in self-management of RA (web app)[t]	Self-efficacy (ASES)[u]	PI[v]: 83.9 (19.0); 9 mths PI: 84.1 (16.3)	PI: 68.5 (23.8); 9 mths PI: 68.6 (23.3)	PI: <.001; 9 mths PI: <.001
[33]	96	21 wks	Activity tracking with pedometer w[w] or wo[x] step targets	Physical activity (steps/day); Fatigue (PROMIS[y]-fatigue)	Steps/day w: 1656 (2161), wo: 1441 (2829); Fatigue w: –4.8 (7.7), wo: –3.2 (7.2)	Steps/day: –747 (3064); Fatigue: –1.6 (8.1)	Steps/day: .003; Fatigue: .21

[a]Ref: reference.

[b]Pts: patients.

[c]RA: rheumatoid arthritis.

[d]In the results column, between-group differences are presented.

[e]mths: months.

[f]P-SEMS: Patient-Reported Outcomes Measurement Information System Self-Efficacy Managing Symptoms.

[g]PAM: Patient Activation Measure.

[h]SD unknown.

[i]Intention-to-treat analysis.

[j]CQR-9: 9-item Compliance Questionnaire-Rheumatology.

[k]wks: weeks.

[l]—: not available.

[m]Unstandardized beta-coefficient (β) of multilevel linear model, including time exposed to intervention; no univariate results presented.

[n]POMI: Prescription Opioid Misuse Index.

[o]No significant differences were found in medication overuse.

[p]No significant differences in the other groups.

[q]N/A: not applicable.

[r]PEPPI-5: 5-item Perceived Efficacy in Patient-Physician Interactions.

[s]CDAI: Clinical Disease Activity Index.

[t]No primary outcome was defined.

[u]ASES: Arthritis Self-Efficacy Scale.

[v]PI: postintervention.

[w]w: with.

[x]wo: without.

[y]PROMIS: Patient-Reported Outcomes Measurement Information System.

SMS Reminders

Three RCTs investigated the use of SMS reminders. Mary et al [26] evaluated the impact of weekly text messages on medication adherence in patients taking methotrexate for RA. Study patients who received reminder text messages on the day they had to take their methotrexate had a greater increase in medication adherence, as assessed using the 9-item Compliance Questionnaire-Rheumatology (CQR-9), compared with a control group and a group receiving a 15-minute pharmacist counselling session [26].

Two other RCTs [27,28] evaluated the effect of three individual motivational counselling sessions and individually tailored text message reminders on reducing daily sitting time. Initially, the pilot study [27] showed feasibility of the study design and acceptability of the intervention; a second study [28] evaluated the effect of the intervention on a larger population. In the intervention group of both studies, individually tailored text messages were sent to each participant to remind them of their behavioral goal(s). Participants could indicate their desired frequency of reminders (between 1 and 5 per week) [27,28]. Ultimately, patients in the intervention group reduced their daily sitting time by 1.61 hours per day [28].

Web Apps

Five RCTs investigated the use of an online platform [13,29-32]. The online platforms were used as informative tools for patients and offered a way to self-monitor disease activity. The platforms demonstrated statistically significant effects in terms of self-management skills, patient empowerment, patient-physician interaction, and physical activity [13,29,30,32].

Allam et al [29] evaluated the effect of different sections of a web app on physical activity, health care utilization, and medication overuse. In the 4 intervention arms of the trial, patients received access to (1) the information section of the web app alone, (2) the information section combined with the social support section of the app, (3) the gaming section of the app, or (4) both the social support section and the gaming section. The intervention arms were compared with a control group that received no access to the web app. Patients that had access to the social support sections on the website decreased health care utilization and medication overuse, and patients with access to gamification features alone or combined with social support increased physical activity and decreased health care utilization [29].

Shigaki et al [32] evaluated the use of an online platform to improve self-efficacy, quality of life, health status, and pain. The platform combined individual and community features. Individual features included educational modules encouraging positive coping strategies for enhancing self-efficacy. In addition to online features, each member was provided with one-on-one leader support through weekly phone contact, typically lasting between 15 and 30 minutes. The platform improved self-efficacy and quality of life in the intervention group: no statistically significant improvements were seen in terms of health status or pain in the intervention group. Data collected through self-monitoring with patient-reported outcomes (PROs) were used for clinical decision making in one study [31]. In the study, PROs were remotely collected to evaluate disease activity by making use of a web app. A total of 44 patients were randomly allocated into 2 groups: the telemonitoring intensive strategy (TIS) group or the conventional strategy (control) group. In the TIS group, patients were monitored intensively and treated according to strict protocols. More patients in the TIS group achieved Clinical Disease Activity Index (CDAI) remission versus patients in the control group (38.1% versus 25% at 1 year; $P<.01$). Moreover, remission was achieved more rapidly, with a median of 20 weeks versus a median of over 36 weeks ($P<.001$)) [31].

Smartphone App

A smartphone app was used in one study [8]. In the study, the use and feasibility of optical imaging through the smartphone were tested to monitor the progression of RA inflammation and deformity in patients' hands. Inflammation and deformity were recorded by taking photos using the smartphone's camera with a standardized procedure. The app also supported self-management behaviors with features to monitor symptoms and record lifestyle and environmental data (eg, diet, activity, and weather). After 6 months of app use, there was a statistically significant improvement in Patient-Reported Outcomes Measurement Information System (PROMIS) Self-Efficacy Managing Symptoms (P-SEMS) [8]. Results of the accuracy of the use of optical imaging for diagnosing flares of arthritis were not presented [8].

Activity Tracking

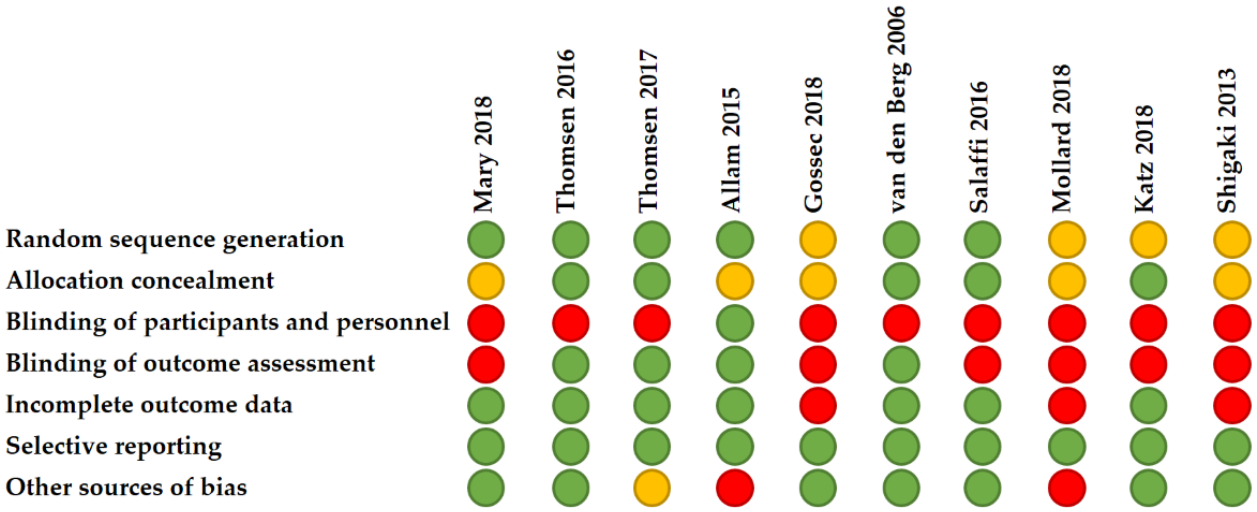
A pedometer was used in one study [33], as a tool to monitor and improve physical activity. The overarching aim of the study was to reduce fatigue, measured with the 7-item PROMIS-fatigue questionnaire. Patients were randomly assigned into one of three parallel arms: (1) physical activity education only, (2) a pedometer plus step-monitoring diary, or (3) a pedometer that visualized step counts combined with goal setting. In both pedometer groups, the number of steps increased significantly, and changes differed significantly with the education-only group. Within-group changes in PROMIS-fatigue

scores in both intervention groups were statistically significant (pedometer plus diary group, *P*=.02, and pedometer plus goal setting group, *P*<.001) [33]. However, the between-group difference in fatigue scores over time between the intervention and control groups were not statistically significant (*P*=.21) [33].

Risk-of-Bias Assessment

At least one domain of the risk-of-bias tool was scored as "at risk for bias in all studies." Blinding participants was not performed in 9 out of 10 studies. No studies were classified as at high risk for selective reporting. The majority of studies (6/10, 60%) were at risk for bias on multiple domains, and in 70% (7/10) of studies, there was an unclear risk for bias on at least one domain. Results are presented in Figure 2.

Figure 2. Assessment of risk of bias with the Cochrane Collaboration's tool. Green=low risk of bias, red=high risk of bias, and orange=unclear risk of bias.



Discussion

Summary

With this review, we provide an overview of the effects of mHealth interventions tested in clinical trials in patients with RA. We identified 10 studies that examined 4 different types of mHealth tools [8,13,26-33]. Web apps were the most tested mHealth intervention [13,29,30,32]. In 9 studies, significant desirable effects were reported on five main outcome measures: medication compliance (SMS reminders) [31], various domains of physical activity (pedometer, SMS reminders, and web apps) [13,27-29,33], percentage of patients that reach remission after 1 year (web app) [31], patient-physician interaction (web app) [30], and self-efficacy (smartphone and web apps) [8,32]. However, all studies were at risk for bias on at least one domain. Due to the heterogeneity in study outcomes and methods used, and the risk of bias in all studies, the promising effect of asynchronous mHealth on all the aforementioned outcomes needs to be corroborated in future studies.

Principal Results

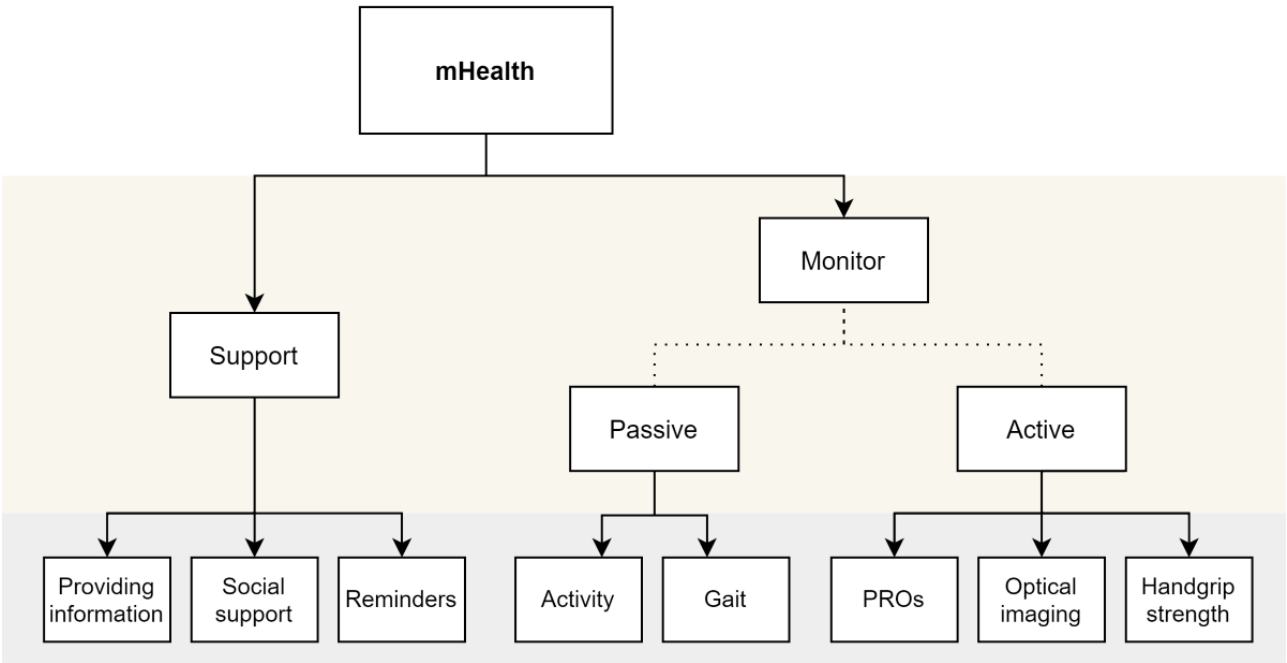
The results of the studies in this review show that it is possible with mHealth interventions to effectively monitor patients to achieve remission sooner [31]. Also, by sending reminders, mHealth tools can motivate patients to improve medication compliance [26], as well as to be more physically active [13,28,33]. Furthermore, mHealth can improve the self-efficacy of patients with RA [8,30,32]. If these results are corroborated, mHealth may contribute to better overall health [34,35], less health care usage, and possibly lower health care costs. In addition, we hypothesize that the value of patient-physician consultations that do happen can be increased, as several reviewed studies reported higher self-efficacy scores and better patient-physician communication. However, outcomes and study methods used in the mHealth studies were very heterogeneous, and studies were often at risk of methodological biases, which was also found in other mHealth reviews [15,36]. Heterogeneity was found in the wide range of interventions (smartphone apps, web apps, pedometers, SMS reminders) and chosen outcome measures. For instance, several different outcomes were measured (eg, self-efficacy, physical activity, and time to reach remission) and several different measures of self-efficacy were used (eg, Arthritis Self-Efficacy Scale, Patient Activation Measure, and P-SEMS) [8,32]. Furthermore, bias could have been introduced into most of the study results, as blinding of patients to their intervention was rarely achieved. Even though practical limitations often impede blinding in mHealth studies, it is recommended to ensure comparable experiences between control and intervention participants [37]. The study of Allam et al [29] illustrates that this can be achieved by utilizing a factorial study design; other options include an early-versus-delayed study design or partial blinding [37].

Expectations

Based on the limited evidence available from mHealth intervention studies, we will cautiously discuss some potentially promising implementations and expectations. We distinguished two main implementations of asynchronous mHealth: (1) monitoring of patients, which can be subdivided into active and passive monitoring, and (2) supporting patients in their health behavior (Figure 3). With the help of remote monitoring through mHealth, unnecessary consultations might be prevented. In inflammatory bowel disease, a monitoring mHealth intervention

has shown to reduce health care visits (33%) and hospital admissions, without increasing disease activity or decreasing patient satisfaction [19]. We expect that monitoring mHealth tools could also be very useful to patients with a stable, chronic rheumatic disease. Often, the value of a consultation is low, as 75% of patients with RA in routine clinical follow-up have low disease activity or are in remission [38]. With monitoring mHealth tools, these consultations could be avoided, which could in turn decrease health care costs. Fewer consultations, however, may also delay biological tapering or treatment when necessary. The safety and cost-effectiveness of this form of monitoring mHealth is currently investigated in two RCTs [20,39]. We anticipate that the use of monitoring mHealth interventions (telemonitoring) will increase based on the increasing number of studies performed with mHealth and the increasing use of smartphones in the population. For patients with RA and their health care providers, objective parameters of disease activity that do not require patient effort (ie, filling in a questionnaire) but are automatically collected would be ideal. That is because passive, remote monitoring eliminates the need for active sharing of disease outcomes and would therefore (partly) surpass adherence issues. Ultimately, this would lead to less missing data and lower the burden for patients. Devices can already collect and share these data without any participation of the patient, and uses in rheumatology are being explored [10,11,40,41] (Figure 3). mHealth that supports patients in their health behavior (supporting mHealth) can help patients adhere better to healthy lifestyles and medication, and can play a significant role in helping patients to cope with their disease by means of social support, education, and improving self-efficacy [13,26-30,32]. In this review, studies used supportive mHealth to get patients more physically active, improve their self-efficacy, and increase their medication adherence [13,26-30,32]. In other medical fields, similar results were found with positive effects in terms of lifestyle and medication adherence [42-44]. This indicates that supportive mHealth may play an important role in preventive (health behavior change) medicine. However, one important gap in our knowledge here is that it is unclear how long the effects of these interventions last. No study in this review had a follow-up of longer than 1 year, and effects often decreased over time. This is likely due to the decrease in adherence to the intervention, which impedes the long-term impact. To increase adherence, some studies have reported on the use of persuasive elements in mHealth tools, such as the use of gamification or persuasive principles in text reminders [29,45]. We expect this to be an important line of research to increase adherence to mHealth tools and to ultimately optimize the impact of mHealth interventions.

Figure 3. Identification of asynchronous mobile health (mHealth) uses in rheumatoid arthritis. PROs: patient-reported outcomes.



Strengths and Limitations

This review was broad in scope and did not focus on one type of mHealth intervention, allowing it to provide a clear overview of the current position of all mHealth tools tested in clinical trials in patients with RA. To maintain a broad scope, the study regarded web apps as mHealth tools because web apps are easily accessible through mobile phones. However, it is possible, that the adoption of, or adherence to, web apps is different on computers and mobile phones. This might lead to other outcomes of the interventions, which should be further examined in future research. Lastly, it is possible that due to biased preference, only studies with positive results are published, which could have led to an overrepresentation of positive effects in this review. However, there is little evidence that this occurred, as we examined trial registries (clinicaltrials.gov and isrctn.com) and encountered only one trial that was completed more than 2 years ago, which had not published its results.

Conclusion

There is a limited number of studies assessing the effect of asynchronous mHealth interventions in patients with RA. The available studies show that asynchronous mHealth interventions have been used to monitor patients or to support them in their health behavior. The reviewed studies reported significant beneficial results of SMS reminders, web apps, and smartphone apps on several outcomes. However, study methods varied widely, all studies were at risk of bias, and follow-up length was often short. Therefore, the results of the review indicate that all reviewed types of mHealth interventions show some promise, but also that these results need to be corroborated in future studies.

Authors' Contributions

BS and WB wrote and edited the manuscript, BS and WB screened the abstracts, and BS and PdB assessed the quality of the studies. All other authors provided feedback and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 17 KB - mhealth_v8i11e19260_app1.docx]

References

1. Myasoedova E, Crowson CS, Kremers HM, Therneau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising?: results from Olmsted County, Minnesota, 1955-2007. Arthritis Rheum 2010 Jun 26;62(6):1576-1582 [FREE Full text] [doi: 10.1002/art.27425] [Medline: 20191579]
2. Ward IM, Schmidt TW, Lappan C, Battafarano DF. How Critical is Tele-Medicine to the Rheumatology Workforce? Arthritis Care Res (Hoboken) 2016 Oct 19;68(10):1387-1389 [FREE Full text] [doi: 10.1002/acr.22853] [Medline: 26866514]
3. World Health Organisation. World Report on Ageing and Health. 2015. URL: https://www.who.int/ageing/events/world-report-2015-launch/en/ [accessed 2019-10-01]
4. Smolen J, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis 2016 Jan;75(1):3-15. [doi: 10.1136/annrheumdis-2015-207524] [Medline: 25969430]
5. Sheikh A, Sood HS, Bates DW. Leveraging health information technology to achieve the "triple aim" of healthcare reform. J Am Med Inform Assoc 2015 Jul 15;22(4):849-856 [FREE Full text] [doi: 10.1093/jamia/ocv022] [Medline: 25882032]
6. van Riel P, Alten R, Combe B, Abdulganieva D, Bousquet P, Courtenay M, et al. Improving inflammatory arthritis management through tighter monitoring of patients and the use of innovative electronic tools. RMD Open 2016 Nov 24;2(2):e000302. [doi: 10.1136/rmdopen-2016-000302] [Medline: 27933206]
7. World Health Organization. New horizons for health through mobile technologies: second global survey on eHealth. 2011. URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2019-10-01]
8. Mollard E, Michaud K. A Mobile App With Optical Imaging for the Self-Management of Hand Rheumatoid Arthritis: Pilot Study. JMIR Mhealth Uhealth 2018 Oct 29;6(10):e12221 [FREE Full text] [doi: 10.2196/12221] [Medline: 30373732]
9. Venkatesh, Thong, Xu. Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. MIS Quarterly 2012;36(1):157. [doi: 10.2307/41410412]
10. Yamada M, Aoyama T, Mori S, Nishiguchi S, Okamoto K, Ito T, et al. Objective assessment of abnormal gait in patients with rheumatoid arthritis using a smartphone. Rheumatol Int 2012 Dec 23;32(12):3869-3874. [doi: 10.1007/s00296-011-2283-2] [Medline: 22193221]
11. Nishiguchi S, Ito H, Yamada M, Yoshitomi H, Furu M, Ito T, et al. Self-assessment tool of disease activity of rheumatoid arthritis by using a smartphone application. Telemed J E Health 2014 Mar;20(3):235-240. [doi: 10.1089/tmj.2013.0162] [Medline: 24404820]
12. Espinoza F, Le Blay P, Coulon D, Lieu S, Munro J, Jorgensen C, et al. Handgrip strength measured by a dynamometer connected to a smartphone: a new applied health technology solution for the self-assessment of rheumatoid arthritis disease activity. Rheumatology (Oxford) 2016 May 10;55(5):897-901. [doi: 10.1093/rheumatology/kew006] [Medline: 26867731]
13. van den Berg MH, Ronday HK, Peeters AJ, le Cessie S, van der Giesen FJ, Breedveld FC, et al. Using internet technology to deliver a home-based physical activity intervention for patients with rheumatoid arthritis: A randomized controlled trial. Arthritis Rheum 2006 Dec 15;55(6):935-945 [FREE Full text] [doi: 10.1002/art.22339] [Medline: 17139640]
14. Mechanic OJ, Persaud Y, Kimball AB. Telehealth Systems. StatPearls (Internet). Treasure Island: StatPearls Publishing; 2020. URL: https://www.ncbi.nlm.nih.gov/books/NBK459384/

15. Piga M, Cangemi I, Mathieu A, Cauli A. Telemedicine for patients with rheumatic diseases: Systematic review and proposal for research agenda. Semin Arthritis Rheum 2017 Aug;47(1):121-128. [doi: 10.1016/j.semarthrit.2017.03.014] [Medline: 28420491]
16. McDougall JA, Ferucci ED, Glover J, Fraenkel L. Telerheumatology: A Systematic Review. Arthritis Care Res (Hoboken) 2017 Oct 22;69(10):1546-1557 [FREE Full text] [doi: 10.1002/acr.23153] [Medline: 27863164]
17. Roberts LJ, Lamont EG, Lim I, Sabesan S, Barrett C. Telerheumatology: an idea whose time has come. Intern Med J 2012 Oct 10;42(10):1072-1078. [doi: 10.1111/j.1445-5994.2012.02931.x] [Medline: 22931307]
18. Deshpande A, Khoja S, Lorca J, McKibbon A, Rizo C, Husereau D, et al. Asynchronous telehealth: a scoping review of analytic studies. Open Med 2009 Jun 02;3(2):e69-e91 [FREE Full text] [Medline: 19946396]
19. de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, Becx MC, Maljaars JP, Cilissen M, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): a pragmatic, multicentre, randomised controlled trial. The Lancet 2017 Sep;390(10098):959-968. [doi: 10.1016/s0140-6736(17)31327-2]
20. Seppen BF, L'ami MJ, Duarte Dos Santos Rico S, Ter Wee MM, Turkstra F, Roorda LD, et al. A Smartphone App for Self-Monitoring of Rheumatoid Arthritis Disease Activity to Assist Patient-Initiated Care: Protocol for a Randomized Controlled Trial. JMIR Res Protoc 2020 Feb 19;9(2):e15105 [FREE Full text] [doi: 10.2196/15105] [Medline: 32130182]
21. Seppen BF, Wiegel J, L'ami MJ, Duarte Dos Santos Rico S, Catarinella FS, Turkstra F, et al. Feasibility of Self-Monitoring Rheumatoid Arthritis With a Smartphone App: Results of Two Mixed-Methods Pilot Studies. JMIR Form Res 2020 Sep 21;4(9):e20165 [FREE Full text] [doi: 10.2196/20165] [Medline: 32955447]
22. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143 [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009 Oct;62(10):1006-1012. [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]
24. 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]
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(oct18 2):d5928-d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
26. Mary A, Boursier A, Desailly Henry I, Grados F, Séjourné A, Salomon S, et al. Mobile Phone Text Messages and Effect on Treatment Adherence in Patients Taking Methotrexate for Rheumatoid Arthritis: A Randomized Pilot Study. Arthritis Care Res (Hoboken) 2019 Oct 27;71(10):1344-1352 [FREE Full text] [doi: 10.1002/acr.23750] [Medline: 30192070]
27. Thomsen T, Aadahl M, Beyer N, Hetland ML, Løppenthin K, Midtgaard J, et al. Motivational counselling and SMS-reminders for reduction of daily sitting time in patients with rheumatoid arthritis: a descriptive randomised controlled feasibility study. BMC Musculoskelet Disord 2016 Oct 18;17(1):434 [FREE Full text] [doi: 10.1186/s12891-016-1266-6] [Medline: 27756265]
28. Thomsen T, Aadahl M, Beyer N, Hetland ML, Løppenthin K, Midtgaard J, et al. The efficacy of motivational counselling and SMS reminders on daily sitting time in patients with rheumatoid arthritis: a randomised controlled trial. Ann Rheum Dis 2017 Sep 05;76(9):1603-1606. [doi: 10.1136/annrheumdis-2016-210953] [Medline: 28584189]
29. Allam A, Kostova Z, Nakamoto K, Schulz PJ. The effect of social support features and gamification on a Web-based intervention for rheumatoid arthritis patients: randomized controlled trial. J Med Internet Res 2015 Jan 09;17(1):e14 [FREE Full text] [doi: 10.2196/jmir.3510] [Medline: 25574939]
30. Gossec L, Cantagrel A, Soubrier M, Berthelot J, Joubert J, Combe B, et al. An e-health interactive self-assessment website (Sanoia) in rheumatoid arthritis. A 12-month randomized controlled trial in 320 patients. Joint Bone Spine 2018 Dec;85(6):709-714. [doi: 10.1016/j.jbspin.2017.11.015] [Medline: 29246532]
31. Salaffi F, Carotti M, Ciapetti A, Di Carlo M, Gasparini S, Farah S, et al. Effectiveness of a telemonitoring intensive strategy in early rheumatoid arthritis: comparison with the conventional management approach. BMC Musculoskelet Disord 2016 Apr 2;17(1). [doi: 10.1186/s12891-016-1002-2]
32. Shigaki CL, Smarr KL, Siva C, Ge B, Musser D, Johnson R. RAHelp: an online intervention for individuals with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2013 Oct;65(10):1573-1581 [FREE Full text] [doi: 10.1002/acr.22042] [Medline: 23666599]
33. Katz P, Margaretten M, Gregorich S, Trupin L. Physical Activity to Reduce Fatigue in Rheumatoid Arthritis: A Randomized Controlled Trial. Arthritis Care Res (Hoboken) 2018 Jan;70(1):1-10 [FREE Full text] [doi: 10.1002/acr.23230] [Medline: 28378441]
34. Li L, Cui Y, Yin R, Chen S, Zhao Q, Chen H, et al. Medication adherence has an impact on disease activity in rheumatoid arthritis: a systematic review and meta-analysis. PPA 2017 Aug;Volume 11:1343-1356. [doi: 10.2147/ppa.s140457]
35. Verhoeven F, Tordi N, Prati C, Demougeot C, Mougin F, Wendling D. Physical activity in patients with rheumatoid arthritis. Joint Bone Spine 2016 May;83(3):265-270. [doi: 10.1016/j.jbspin.2015.10.002] [Medline: 26639220]

36. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
37. Shcherbatykh I, Holbrook A, Thabane L, Dolovich L. Methodologic Issues in Health Informatics Trials: The Complexities of Complex Interventions. Journal of the American Medical Informatics Association 2008 Sep 01;15(5):575-580. [doi: 10.1197/jamia.m2518]
38. Haugeberg G, Hansen IJW, Soldal DM, Sokka T. Ten years of change in clinical disease status and treatment in rheumatoid arthritis: results based on standardized monitoring of patients in an ordinary outpatient clinic in southern Norway. Arthritis Res Ther 2015 Aug 20;17(1). [doi: 10.1186/s13075-015-0716-0]
39. Wang P, Luo D, Lu F, Elias JS, Landman AB, Michaud KD, et al. A Novel Mobile App and Population Management System to Manage Rheumatoid Arthritis Flares: Protocol for a Randomized Controlled Trial. JMIR Res Protoc 2018 Apr 11;7(4):e84 [FREE Full text] [doi: 10.2196/resprot.8771] [Medline: 29643053]
40. Gossec L, Guyard F, Leroy D, Lafargue T, Seiler M, Jacquemin C, et al. Detection of Flares by Decrease in Physical Activity, Collected Using Wearable Activity Trackers in Rheumatoid Arthritis or Axial Spondyloarthritis: An Application of Machine Learning Analyses in Rheumatology. Arthritis Care Res (Hoboken) 2019 Oct 27;71(10):1336-1343 [FREE Full text] [doi: 10.1002/acr.23768] [Medline: 30242992]
41. Jacquemin C, Servy H, Molto A, Sellam J, Foltz V, Gandjbakhch F, et al. Physical Activity Assessment Using an Activity Tracker in Patients with Rheumatoid Arthritis and Axial Spondyloarthritis: Prospective Observational Study. JMIR Mhealth Uhealth 2018 Jan 02;6(1):e1 [FREE Full text] [doi: 10.2196/mhealth.7948] [Medline: 29295810]
42. Kosse RC, Bouvy ML, de Vries TW, Koster ES. Effect of a mHealth intervention on adherence in adolescents with asthma: A randomized controlled trial. Respir Med 2019 Mar;149:45-51. [doi: 10.1016/j.rmed.2019.02.009] [Medline: 30803885]
43. Wu X, Guo X, Zhang Z. The Efficacy of Mobile Phone Apps for Lifestyle Modification in Diabetes: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
44. Jeminiwa R, Hohmann L, Qian J, Garza K, Hansen R, Fox BI. Impact of eHealth on medication adherence among patients with asthma: A systematic review and meta-analysis. Respir Med 2019 Mar;149:59-68. [doi: 10.1016/j.rmed.2019.02.011] [Medline: 30803887]
45. Hallsworth M, Berry D, Sanders M, Sallis A, King D, Vlaev I, et al. Stating Appointment Costs in SMS Reminders Reduces Missed Hospital Appointments: Findings from Two Randomised Controlled Trials. PLoS One 2015 Sep 14;10(9):e0137306 [FREE Full text] [doi: 10.1371/journal.pone.0137306] [Medline: 26366885]

Abbreviations

CDAI: Clinical Disease Activity Index
CQR-9: 9-item Compliance Questionnaire-Rheumatology
mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PROs: patient-reported outcomes
PROMIS: Patient-Reported Outcomes Measurement Information System
P-SEMS: Patient-Reported Outcomes Measurement Information System Self-Efficacy Managing Symptoms
RA: rheumatoid arthritis
RCT: randomized controlled trial
TIS: telemonitoring intensive strategy

Edited by G Eysenbach; submitted 10.04.20; peer-reviewed by H Vonkeman, M Piga, S Guo, T Pelle; comments to author 29.06.20; revised version received 18.08.20; accepted 21.08.20; published 05.11.20.

Please cite as:
Seppen BF, den Boer P, Wiegel J, ter Wee MM, van der Leeden M, de Vries R, van der Esch M, Bos WH
Asynchronous mHealth Interventions in Rheumatoid Arthritis: Systematic Scoping Review
JMIR Mhealth Uhealth 2020;8(11):e19260
URL: http://mhealth.jmir.org/2020/11/e19260/
doi:10.2196/19260
PMID:33151161

Original Paper

Real-World Evidence of User Engagement With Mobile Health for Diabetes Management: Longitudinal Observational Study

Anna-Katharina Böhm[1], MSc; Morten Lind Jensen[2], DPhil, DPM, MD; Mads Reinholdt Sørensen[2], DPhil; Tom Stargardt[1], Prof Dr

[1]Hamburg Center for Health Economics, University of Hamburg, Hamburg, Germany
[2]Novo Nordisk A/S, Søborg, Denmark

Corresponding Author:
Anna-Katharina Böhm, MSc
Hamburg Center for Health Economics
University of Hamburg
Esplanade 36
Hamburg, 20354
Germany
Phone: 49 40428381627
Email: Anna-Katharina.Boehm@uni-hamburg.de

Abstract

Background: Patient support apps have risen in popularity and provide novel opportunities for self-management of diabetes. Such apps offer patients to play an active role in monitoring their condition, thereby increasing their own treatment responsibility. Although many health apps require active user engagement to be effective, there is little evidence exploring engagement with mobile health (mHealth).

Objective: This study aims to analyze the extent to which users engage with mHealth for diabetes and identify patient characteristics that are associated with engagement.

Methods: The analysis is based on real-world data obtained by Novo Nordisk's Cornerstones4Care Powered by Glooko diabetes support app. User engagement was assessed as the number of active days and using measures expressing the persistence, longevity, and regularity of interaction within the first 180 days of use. Beta regressions were estimated to assess the associations between user characteristics and engagement outcomes for each module of the app.

Results: A total of 9051 individuals initiated use after registration and could be observed for 180 days. Among these, 55.39% (5013/9051) used the app for one specific purpose. The average user activity ratio varied from 0.05 (medication and food) to 0.55 (continuous glucose monitoring), depending on the module of the app. Average user engagement was lower if modules required manual data entries, although the initial uptake was higher for these modules. Regression analyses further revealed that although more women used the app (2075/3649, 56.86%), they engaged significantly less with it. Older people and users who were recently diagnosed tended to use the app more actively.

Conclusions: Strategies to increase or sustain the use of apps and availability of health data may target the mode of data collection and content design and should take into account privacy concerns of the users at the same time. Users' engagement was determined by various user characteristics, indicating that particular patient groups should be targeted or assisted when integrating apps into the self-management of their disease.

(JMIR Mhealth Uhealth 2020;8(11):e22212) doi:10.2196/22212

KEYWORDS

user engagement; user activity; mHealth; diabetes mellitus; diabetes apps

Introduction

Background and Aims

Diabetes causes elevated blood glucose (BG) because the body is not able to produce any or sufficient amounts of insulin or is not capable of processing it effectively. Patients have an increased risk of developing life-threatening complications caused by prolonged elevated glucose levels [1]. Although for people with early stage type 2 diabetes (T2D), a healthy diet, exercise, and weight loss may be sufficient to maintain near normal BG, people with type 1 diabetes (T1D) or progressed T2D require insulin therapy and more intensive glucose monitoring [2]. Thus, if treated adequately, complications can be avoided or delayed, but this requires thorough self-care and adherence to treatment [3].

The ongoing digital revolution combined with the increasing prevalence of diabetes has led to an abundance of mobile health (mHealth) apps promising novel opportunities for self-monitoring and treatment guidance for diabetes patients [4]. Functions such as providing health information, medication reminders, remote monitoring, and mobile analytics may increase the users' health literacy, support patients to play a more active role in managing their disease, and promote adherence to treatment [4,5]. In addition, mHealth may enable sharing of data between patients and health care professionals (HCPs), which could support an improved patient-HCP dialogue [6]. Thus, if utilized optimally, mHealth may be an opportunity to increase quality of care while at the same time offering the potential to reduce costs for health care systems [7-9].

On the other hand, factors such as low digital competence or limited availability of technology could reduce the use and cost-effectiveness of mHealth. In addition, data privacy can be a concern for patients and may lead users to omit use or enter artificial information [10,11]. Furthermore, it is inherently difficult for patients and HCPs to select the most appropriate app for the individual patient, as there are over 300,000 health apps with different key features available on the market [12].

Although the variety and popularity of mHealth for diabetes patients is increasing, there is limited evidence about who uses the apps, the users' engagement, and the apps' effectiveness [11]. Surveys from the United States and Germany reveal that users of health apps were younger individuals; more often female; had an advanced education, higher income, and better health literacy [13-16]. However, these studies analyzed health apps in general and ignored disease-specific characteristics of use. Similarly, studies exploring user engagement often focused on other disease areas than diabetes [17-20]. A systematic review analyzing the pattern of user engagement of technology-based interventions for T2D patients emphasized the need for studies reporting on engagement, examining associations between user characteristics and engagement, and standardizing how user engagement is reported [21].

Objectives

This study explores real-world data obtained by Novo Nordisk's Cornerstones4Care (C4C) patient support app (the app) for diabetes. The aim of this study was to study the use of mHealth among diabetes patients, here represented by the users of the app. In particular, the study aims to analyze how intensively users engage with the app and to identify patient characteristics that are associated with engagement with the app. Addressing these aims will help to understand how mHealth apps can be used effectively and what areas could be optimized.

Methods

General Overview of the App

C4C is a US patient support program for diabetes patients, funded by Novo Nordisk A/S. It includes the option to use the *Cornerstones4Care Powered by Glooko* mHealth app. The app has been available since June 2017, and approximately 30,000 users have installed the app by October 2019 and provided consent to use the data for research purposes. Installation does not require prescription by an HCP. The patient support program and the app provide information materials and tools to monitor nutrition, fitness, medication, and glucose levels (the modules). The app allows users to synchronize data from BG meters, continuous glucose monitoring (CGM) devices, insulin pumps, and external fitness and health devices. In addition, the user can view historic data and trend graphs, set reminders for testing their BG or taking medication, and add custom notes. Thus, it brings together different diabetes data relevant for effective disease management. In addition to this, the app allows sharing of data between the patient and the HCP.

Study Outcomes

A major challenge when assessing how users engage with mHealth is the lack of consensus on how to assess user engagement [22]. Therefore, we based our choice of engagement metrics on a theoretical framework defining user engagement of technology as a progress comprising 4 distinct stages: point of engagement, period of sustained engagement, disengagement, and reengagement [23]. We established 6 metrics capturing different aspects of user engagement. Table 1 summarizes the application of the conceptual framework in our study.

To account for censoring, that is, shorter observation periods for users who have downloaded the app more recently, we restricted our sample to users who had started using the app before April 26, 2019, and assessed user activity during the first 180 days after initiation.

Table 1. Conceptual framework underlying user engagement metrics based on the study by O'Brien and Toms, 2013.

Conceptual framework and metric names	Definition
Point of engagement	
Activity delay	Days from consent date to first active day
Period of engagement	
User activity ratio	Number of active days divided by number of potentially active days (here 180)
Longevity	Days from first until last active day (maximum=180)
Recency	Average number of days between entries
Regularity	Coefficient of variation of number of days between entries
Discontinuation	
Persistence	Days until 28 days discontinuation
Dropouts	Illustrated in Figure 1
Reengagement	
Engagement pattern	Illustrated in Figure 2

The core metric is the user activity ratio (UAR). It counts the number of active days (ie, days with at least one entry) of user *i* in module *m*, divided by the number of potentially active days (here 180 days):

In addition to this, we investigated the *activity delay*, which reflects the number of days between the consent date and first active use. Moreover, as proposed by Rahman et al [18], *longevity* captures the time from the initial use until the final entry (maximum 180 days). *Persistence* was expressed as the number of days until the first discontinuation of use within 180 days, where discontinuation was defined as a gap of >28 concurrent inactive days. The size of the permissible gap was based on the 0.9 percentile of all average user gaps. If users did not discontinue use, persistence was set to 180. In accordance with Taki et al [20] and Peterson and Carrabis [24], we also assessed the *recency*, which was defined as the average time between active days for each user. To quantify the *regularity* of activity, we additionally calculated the coefficient of variation of the time between active days, that is, the standard deviation of the time between active days divided by the recency. This number allows a direct comparison between the modules.

Finally, we calculated alternative versions of the UAR with the number of potentially active days as the (1) longevity or (2) persistence.

In addition to engagement, the metrics reflect the period and amount of data collection and, therefore, the availability of rich health information. Kendall correlations for nonlinear data were calculated to assess the redundancy of the metrics. All metrics were calculated for each of the 5 modules and for the overall app.

C4C Data

The C4C data set includes all app entries (manual and automated) collected during the period from app launch (June 28, 2017) to October 21, 2019. The data set is structured according to the 5 main modules of the app, namely, *food intake*, *exercise*, *medication intake*, *BG*, and *CGM*. In addition, it contains basic user information. The app contains no personally identifiable data, and users have provided their consent for data sharing. The data were aggregated for analysis.

User Information

Basic user information covers consent information (eg, date of consent), sociodemographic data (age and gender), diabetes type, bodyweight measures (latest height and weight), and a medication profile (medication type, medication name, and time at which the user indicated treatment initiation). Self-reported age is the only mandatory information to be shared when using the app. If information on height and weight was available, we calculated the user's BMI.

On the basis of the medication profile, we further extracted if a user had registered insulin (yes=1 and no=0), other injectables (yes=1 and no=0), orals (yes=1 and no=0), the total number of registered medications, and the exact number of oral substances. On the basis of the medication name, we also extracted if a user registered fast-acting bolus insulin (yes=1 and no=0) because they represent patients with T1D and progressed T2D, thus patients who may require more intensive monitoring. The resulting treatment regimen information served as a proxy for disease severity.

Furthermore, we combined the information on diabetes type and medication profile to determine if a user was newly diagnosed. T1D patients were considered as recently diagnosed if treatment with insulin started within 180 days before the consent date. T2D patients were considered as recently diagnosed if (1) medication was limited to metformin, which is the typical first-line treatment for people with T2D, and (2) treatment started within 180 days before the date of consent.

Food Intake

Data on food intake include information on registered carbs, fat, protein, and calories. The user can either enter the meal manually, select a meal from a list, or scan the barcode of groceries. The data set stores the time the user indicated that the food was consumed and a timestamp marking the data entry.

Exercise

Exercise data provide a daily summary of the user's activity, including the duration, distance, burned calories, steps, floors, and elevation. The user can either enter an activity manually, pick one from a list, or synchronize with an external health app (as the app must be opened to sync, our engagement metrics can serve as a proxy for user activity in [partly] automated modules). The data set covers the day of the daily summary of the user's activity as well as a timestamp of the data record, that is, the day of manual entry by the user or synchronization from an external app.

Medication Intake

Data about medication intake contain information on the registered medication, including the name, type (oral, insulin, and other injectables), and dosage. The user can pick the medication from a predefined list or make a manual entry. The data cover the time at which the user indicated the medication was taken and a timestamp marking the data entry by the user.

BG Values

The app allows the storage of BG values (mg/dL) obtained by BG meters. A BG meter is a medical device used to determine the concentration of glucose in the blood. It detects the level of sugar in capillary blood taken from a finger prick. The data set contains information on the time of BG measurement and a data entry timestamp (manual entry by the user or synchronized from a BG device via Bluetooth; this function is only available in few BG meters). In addition, users can add tags indicating if a reading occurs during fasting or is taken postprandially. Meal tags can also be synched from the BG meter. If available, we used this information to assess if the user's baseline BG, that is, their first preprandial glucose measurement or their first value after nighttime (5:00 AM until 9:00 AM), was well controlled (yes=1 and no=0). The threshold of 80-130 mg/dL was adopted from official recommendation by the American Diabetes Association (ADA) [25].

CGM Data

In contrast to BG meters, which only capture BG at a certain point in time when a user performs the relevant actions, CGM devices typically generate 5-min interval data. An electrode is placed under the skin, while the transmitter sends data to a separate receiver, and the data are then sent to the app via Bluetooth. The data set covers the glucose values (mg/dL) as well as a trend arrow representing the visual change in the glucose trend. It contains information on the measurement time and a timestamp of the phone transmission (similar to the exercise module, the engagement metrics can serve as a proxy for user activity in [partly] automated modules because the app needs to be opened to sync). In this study, baseline glycemic control was assessed based on recommended target ranges established by an international expert panel in 2019 [26]. We determined that users were in good glycemic control if they exceeded the recommended 70% of time in range (TIR; 70-180 mg/dL) within the first 14 days of observations in which sufficient data were available.

Statistical Analysis

Exploratory Data Analysis

Descriptive statistics of user characteristics and study outcomes are presented as means, standard deviations, and ranges for continuous variables and as counts and percentages for categorical variables. If the values of self-reported data were implausible, observations were treated as unknown.

Uptake and discontinuation of use for each module are visualized as graphs. Survival curves were created to show the proportion of users who were still active (on that day or on a later day) over time. Event diagrams, moreover, visually summarize the pattern of user activity.

Regression Analysis

Multivariate regression models were used to examine the associations between user characteristics and engagement (UAR) for each module. Specifically, we estimated general linear models with a beta distribution (for each module, the link function was chosen based on the smallest Akaike information criterion and Bayesian information criterion for model selection). By doing so, we account for the fact that engagement is double bounded by the values zero and one (100% user activity). If a user's UAR was equal to 1, it was set to 0.99. Missing data, which is a major concern in self-reported data sets, were treated as a separate category (unknown) for all variables to retain the full number of observations. The equation of interest can be described as follows:

where *gender* represents the sex of the users, *age group* is a factor variable grouping users according to their age, *bmi group* represents the user's BMI according to the official classification by the World Health Organization (WHO), *diagnosis* reflects the user's diabetes type, *recent* is a proxy indicating if a user was recently diagnosed with diabetes, *insulin* is an indicator specifying if a user is treated with insulin (and therefore more severely ill), *bg control* indicates if the user's fasting BG at baseline is well controlled, and *modules* reflects the total number of actively used modules per user. For a meaningful interpretation, the margins were calculated for each parameter. *P* values less than .10 were considered statistically significant in our analysis.

Sensitivity Analysis

Given that there is no consensus on what constitutes an adequate amount of user engagement, the calculation of activity metrics was repeated based on active weeks rather than active days to confirm the robustness of our findings. In addition, we confirmed the robustness of our findings by applying gamma and logged ordinary least squares regression as alternative regression specifications.

Results

Descriptive Results

User Characteristics

Among the 29,643 users who gave their consent after download, 12,685 initiated usage of the app. Table 2 summarizes user characteristics for the 9051 users meeting our inclusion criteria (ie, they initiated use and could be observed for 180 days); 63.50% (5747/9051) of the users reported being diagnosed with T2D and 13.48% (1220/9051) with T1D. Among users who shared information about their gender, 56.86% (2075/3649) reported to be female. It should be noted that only 40.32% (3649/9051) of all users provided information on their gender. More than half of the users were aged 50 years or older, and the average age of users was 50.4 years. The average BMI was 34.8, that is, in the obesity range according to the official classification by the WHO. However, the number was based on a subset of 36.69% (3321/9051) of users who shared information about their height and weight.

A total of 76.28% (6904/9051) of users set up a medication profile, with an average of 1.9 registered medications. Among these, 63.33% (4372/6904) registered oral medications and 56.34% (3890/6904) registered insulin.

Within users sharing information about their BG, a subset of 90.73% (4550/5015) provided sufficient data to assess their fasting BG, resulting in 34.48% (1569/4550) of well controlled users. Similarly, 86.76% (485/559) of CGM users provided sufficient data to assess their TIR. Among these, 38.14% (185/485) satisfied the recommended 70% of TIR, indicating good glycemic control.

Table 2. Descriptive statistics of user characteristics (N=9051).

Characteristics	Participants	Minimum value of user characteristics	Maximum value of user characteristics	Total unknown[a]
Sociodemographics				
Gender, n (%)				5402
Female	2075 (56.86)	0	1	
Male	1574 (43.14)	0	1	
Age[b] (n=8880), mean (SD)	50.39 (13.48)	16	94	171
<30 years, n (%)	719 (8.10)	0	1	
30-40 years, n (%)	1184 (13.33)	0	1	
40-50 years, n (%)	2103 (23.68)	0	1	
50-60 years, n (%)	2541 (28.61)	0	1	
60-70 years, n (%)	1737 (19.56)	0	1	
>70 years, n (%)	596 (6.71)	0	1	
Body measures				
Height[c] (n=3646), mean (SD)	169.70 (10.57)	142.20	205.70	5405
Weight[d] (n=3507), mean (SD)	100.38 (29.24)	45.36	300.00	5544
BMI[e] (n=3321), mean (SD)	34.83 (9.79)	15.82	106.83	5730
Underweight, n (%)	21 (0.63)	0	1	
Normal weight, n (%)	354 (10.66)	0	1	
Overweight, n (%)	716 (21.56)	0	1	
Obese I, n (%)	851 (25.62)	0	1	
Obese II, n (%)	612 (18.43)	0	1	
Obese III, n (%)	767 (23.10)	0	1	
Diagnosis				
Type, n (%)				894
Type 1 diabetes	1220 (14.96)	0	1	
Type 2 diabetes	5747 (70.45)	0	1	
Prediabetes	715 (8.77)	0	1	
Other	475 (5.82)	0	1	
Duration, n (%)				2562
Newly diagnosed	1780 (27.43)	0	1	
Treatment regimen				2147
Number of registered medications (≥1) (n=6904), mean (SD)	1.86 (0.97)	1	11	
Registered insulin, n (%)	3890 (56.34)	0	1	
Registered bolus insulin, n (%)	2334 (33.81)	0	1	
Registered other injectables, n (%)	1160 (16.80)	0	1	
Registered orals, n (%)	4372 (63.33)	0	1	
Number of oral substances (≥1) (n=4372), mean (SD)	1.38 (0.64)	1	6	
Control of disease, n (%) of well-controlled users				
On the basis of baseline BG[f]	1569 (34.48)	0	1	4501
On the basis of baseline CGM[g]	185 (38.14)	0	1	8566

Characteristics	Participants	Minimum value of user characteristics	Maximum value of user characteristics	Total unknown[a]
Demanded app modules, n (%)				0
Ever used medication tracker	3933 (43.45)	0	1	
Ever used exercise	1579 (17.45)	0	1	
Ever used food	3821 (42.22)	0	1	
Ever used BG	5015 (55.41)	0	1	
Ever used CGM	559 (6.18)	0	1	
User scope				0
Number of modules used (>0) (n=9051) mean (SD)	1.65 (0.83)	0	5	
Used 1 module, n (%)	5013 (55.39)	0	1	
Used 2 modules, n (%)	2465 (27.23)	0	1	
Used 3 modules, n (%)	1349 (14.90)	0	1	
Used 4 modules, n (%)	203 (2.24)	0	1	
Used 5 modules, n (%)	21 (0.23)	0	1	

[a]Total unknown reflects the number of users who did not share their information for each user characteristic.

[b]Age in years (between 16 and 100 years).

[c]Height in cm (between 140 and 220 cm).

[d]Weight in kg (between 45 and 300 kg).

[e]BMI groups according to the definition of the World Health Organization: underweight=BMI<18.5, normal weight=25>BMI≥18.5, overweight=30>BMI≥25, obese I=35>BMI≥30, obese II=40>BMI≥35, obese III=BMI≥40.

[f]BG: blood glucose.

[g]CGM: continuous glucose monitoring.

User Engagement

In addition to the user characteristics, Table 2 summarizes the use of the different modules of the app. Most initiated use of the BG module (5015/9051, 55.41%), medication module (3933/9051, 43.45%), and food module (3821/9051, 42.22%). A total of 17.45% (1579/9051) of all users applied the exercise module and 6.18% (559/9051) used the CGM module. It is noteworthy that more than 50% (5013 of 9051) of the individuals used the app for one single purpose, whereas only a total of 21 users took advantage of the entire app.

Table 3 presents the means and standard deviations of our study outcomes, that is, the 6 different user engagement metrics. Although most initiated use of the BG, medication, and food modules, the UAR was the lowest among the 3 modules (BG 0.07, medication 0.05, and food 0.05). In contrast, the UAR was highest for CGM (0.55) and exercise (0.37), both modules whose initial uptake remained rather low. Accordingly, activity delay was shortest for the medication and food modules (15 and 13 days, respectively), but low persistence (15 and 13 days, respectively) and longevity (23 and 20 days, respectively) showed that a high share of users stopped their engagement with the modules soon after initiation. In contrast, activity delay was longest for the CGM module (83 days), but once initiated, users remained active for an average of 133 days. A similarly high longevity (114 days) was observed for the exercise module. The recency of use was lowest for the exercise module (6 days) and highest for the BG module (18 days). Finally, it should be noted that the UAR turns from lowest to highest for the food module and for the medication module, if the denominator was defined as longevity or persistence, that is, if only the users' active periods were considered.

Table 3. Descriptive statistics of user engagement metrics for each module.

Activity metrics	Medication (n=3933), mean (SD)	Food (n=3821), mean (SD)	Exercise (n=1579), mean (SD)	BG[a] readings (n=5015), mean (SD)	CGM[b] (n=559), mean (SD)	Any (N=9051), mean (SD)
User activity ratio						
Denominator specified as 180 days	0.05 (0.15)	0.05 (0.14)	0.37 (0.35)	0.07 (0.15)	0.55 (0.41)	0.14 (0.27)
Denominator specified as the longevity	0.79 (0.33)	0.82 (0.31)	0.61 (0.32)	0.51 (0.41)	0.72 (0.34)	0.65 (0.39)
Denominator specified as the persistence	0.84 (0.28)	0.86 (0.26)	0.66 (0.29)	0.64 (0.37)	0.83 (0.26)	0.73 (0.33)
Activity delay	14.98 (60.06)	12.73 (54.33)	19.22 (78.10)	20.47 (64.77)	82.55 (149.70)	14.14 (58.96)
Longevity	23.25 (46.95)	19.68 (42.56)	113.7 (78.50)	57.97 (65.32)	133.00 (66.50)	58.56 (71.90)
Persistence	15.20 (36.71)	13.14 (32.45)	100.30 (81.85)	31.11 (49.75)	105.60 (77.49)	40.00 (63.49)
Recency	9.06(18.83)	6.83 (16.76)	5.53 (11.85)	17.87 (26.79)	7.68 (22.47)	12.11 (22.78)
Regularity	0.96 (0.75)	0.92 (0.80)	0.90 (0.54)	0.97 (0.59)	0.96 (1.06)	0.98 (0.71)

[a]BG: blood glucose.

[b]CGM: continuous glucose monitoring.

Figures 1 and 2 support the findings in Table 3. Figure 1 emphasizes that dropout occurred faster for the modules in which use was initiated most often (food, medication, and BG). A large fraction of users dropped out already within the first week, and in the following weeks, the dropout continues to be higher than that for the other 2 modules. In contrast, more than half of the users remained active in the CGM and the exercise modules during the entire 180 days period.

Figure 2 shows the event diagrams visualizing the engagement patterns. Each dotted horizontal line represents a user's interaction with the respective app module. The density of the graph consequently reflects all users' activity within a module. This means that the darker the area under the curve, the more users engaged with the module. Again, the CGM and the exercise modules appear to be used most persistently, whereas the usage of the other 3 modules was often discontinued, as reflected by the white spaces.

Finally, Figure 3 adds information by showing module uptake over time. The general upward trends of the curves indicate that more individuals initiate use than drop out over time, which may indicate an increasing demand for mHealth and self-management of diabetes, for example, through successful marketing activities or other reasons for increasing awareness.

Figure 1. Survival curves visualizing the relative dropout of users per module. Active users are defined as active on that day or on a later day. Day 1 marks the day of initial interaction with each module. BG: blood glucose; CGM: continuous glucose monitoring.

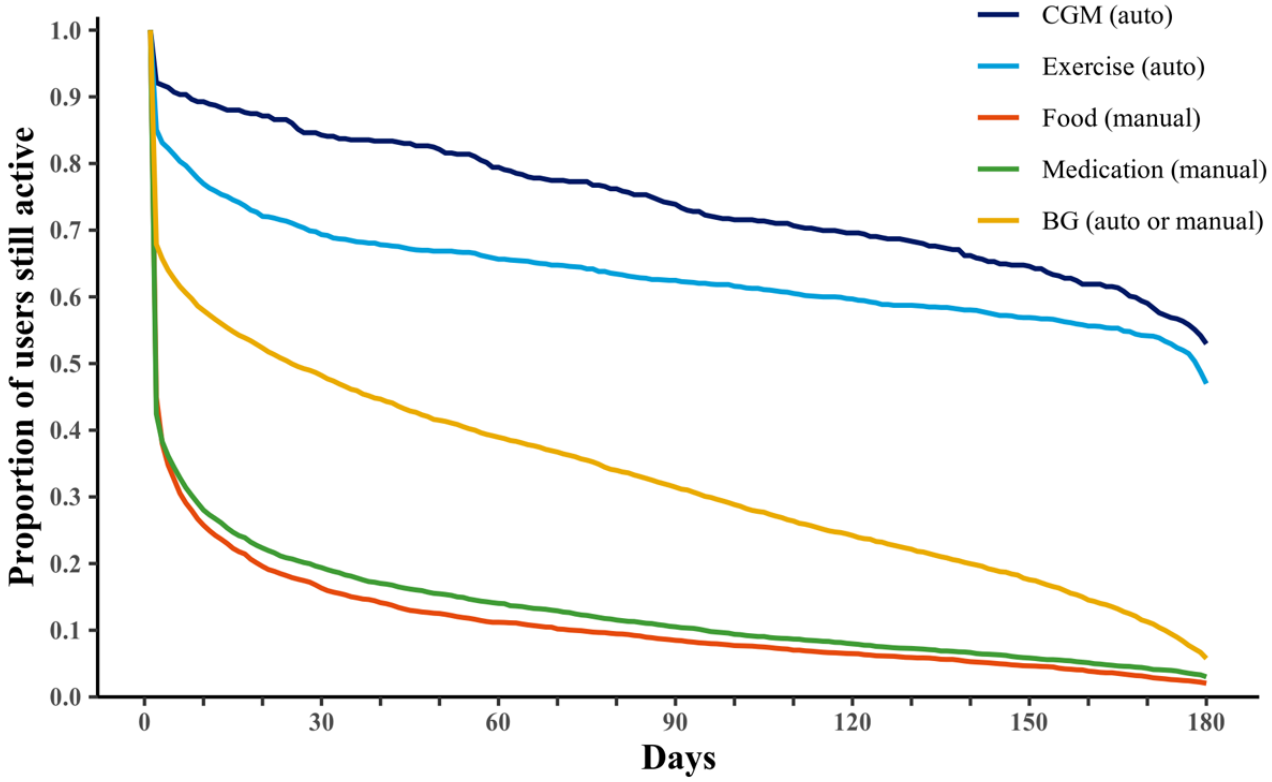


Figure 2. Event diagrams expressing the user activity patterns over time for the 5 modules. A dotted horizontal line represents a user's interaction with the app, and each dot represents an active day. Dots from multiple users and days may be overlapping; however, the density of points shows a pattern of intensity. BG: blood glucose; CGM: continuous glucose monitoring.

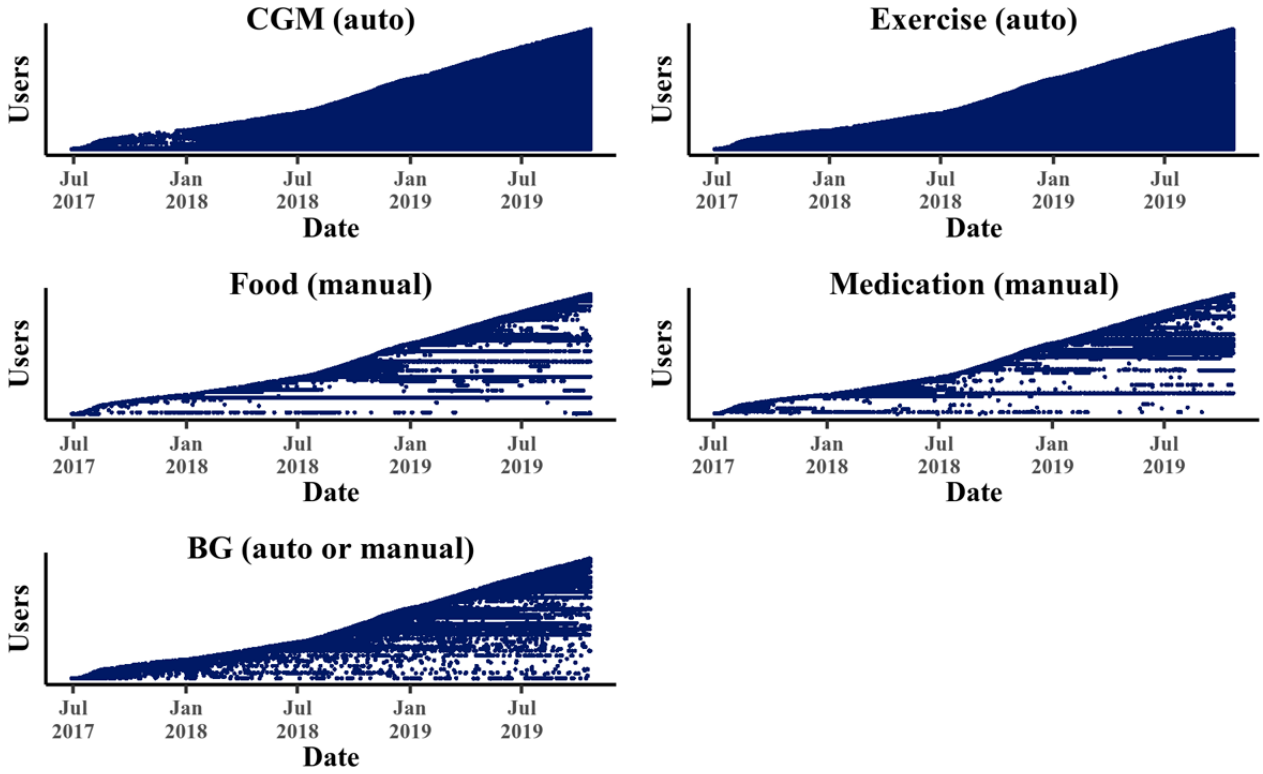
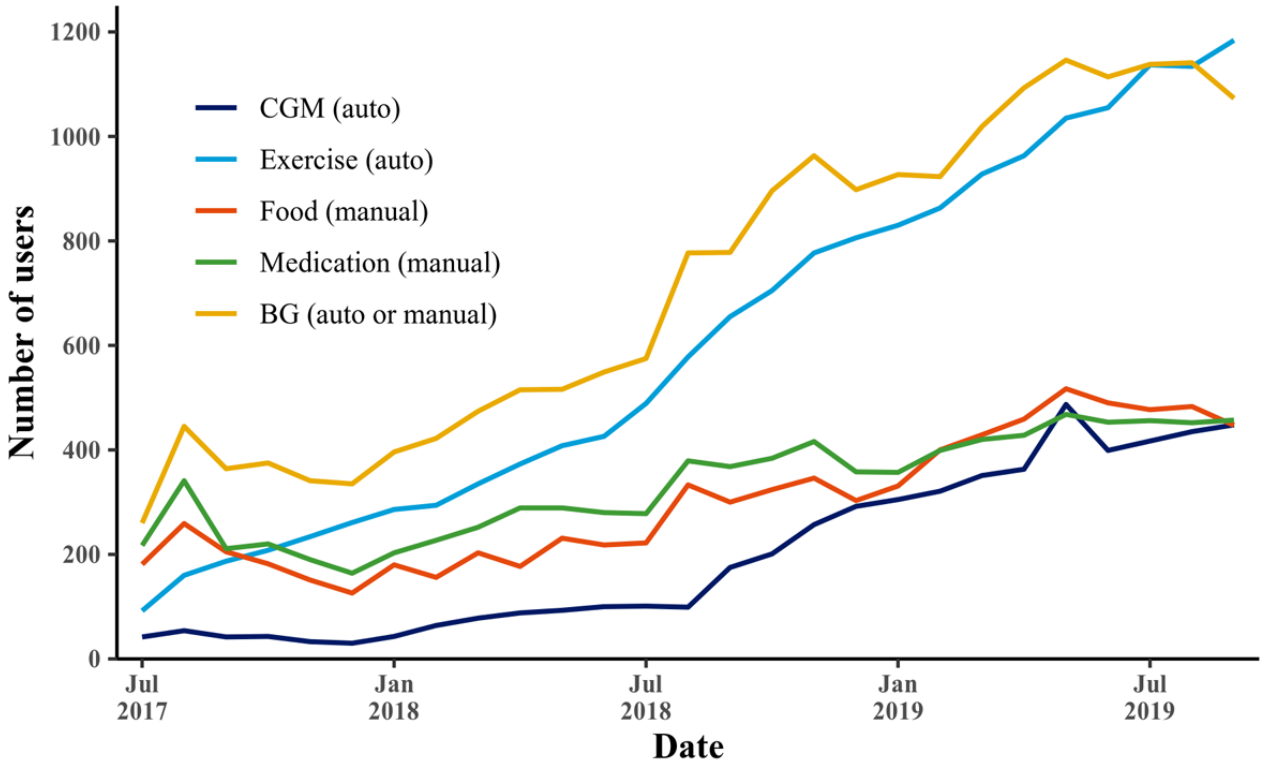


Figure 3. Number of active users per month. The increasing trend over time shows that more users are enrolling than dropping out. BG: blood glucose; CGM: continuous glucose monitoring.



Regression Analysis

Tables 4 and 5 summarize the results of the regression analysis, linking user characteristics and user engagement (UAR). Each column represents one module. For easier interpretation, margins are reported for each parameter and translated into active days.

Although more women used the app (2075/3649, 56.86%), the regression results indicate that they engaged significantly less with the app in 4 of the 5 modules. Only the exercise module was used significantly more by female users. Regression results further show that older users engaged significantly more with the app in all modules except the CGM module. In addition, there is a small tendency that individuals with a more recent first diabetes diagnosis used the app more intensively. The total number of actively used modules was positively associated with the interaction in a single module. Finally, body measure and glycemic control at baseline did not correlate significantly with user intensity in any module. Findings relating to diabetes type and severity of the disease do not point to a clear direction in our analysis.

To confirm robustness, the analysis was repeated with an UAR based on active weeks rather than active days. The results supported the abovementioned findings. When comparing the average number of active days and weeks in each module, it can be concluded that users do not engage with the app on concurrent days but instead interact after breaks of several days.

Applying alternative regression specifications also confirmed the robustness of our findings.

Table 4. Beta regressions estimating the effect of user characteristics on user engagement (user activity ratio) for nonautomated modules.

User characteristics	App module					
	Medication (n=3933; average UAR[a]=9 days)			Food (n=3821; average UAR=9 days)		
	ME[b]	*P* value	Days	ME	*P* value	Days
Sociodemographic						
Gender (reference=male)						
Female	−0.019	<.001	−3.42	−0.014	<.001	−2.52
Unknown	−0.020	<.001	−3.60	−0.015	<.001	−2.70
Age (reference=30-40 years)						
<30 years	0.000	.99	0.00	0.000	.59	0.00
40-50 years	0.003	.23	0.54	0.003	.06	0.54
50-60 years	0.007	.004	1.26	0.007	<.001	1.26
60-70 years	0.009	.003	1.62	0.010	<.001	1.80
>70 years	0.008	.098	1.44	0.009	.03	1.62
Unknown	−0.003	.69	−0.54	0.008	.42	1.44
BMI (reference =normal)						
Underweight (BMI<18.5)	−0.007	.73	−1.26	−0.020	.007	−3.60
Overweight (25≤BMI<30)	0.010	.12	1.80	0.009	.19	1.62
Obese (30≤BMI)	0.003	.50	0.54	−0.001	.81	−0.18
Unknown	0.000	.99	0.00	−0.003	.50	−0.54
Diagnosis						
Type (reference=type 1 diabetes)						
Type 2 diabetes	0.004	.095	0.72	0.000	.89	0.00
Prediabetes	0.016	.097	2.88	0.007	.10	1.26
Other	0.018	.05	3.24	0.003	.48	0.54
Unknown	0.038	<.001	6.84	0.014	.01	2.52
Recently diagnosed (reference=no)						
Yes	0.002	.16	0.36	0.003	.09	0.54
Unknown	−0.012	.16	−2.16	−0.003	.46	−0.54
Disease status						
Treatment (reference=noninsulin)						
Insulin treatment	0.002	.17	0.36	−0.004	.02	−0.72
Unknown	0.014	.12	2.52	0.007	.02	1.26
Blood glucose in control (reference=no)						
Yes	−0.005	.20	−0.90	0.000	.97	0.00
Unknown	−0.015	<.001	−2.70	−0.012	<.001	−2.16
Usage						
Number of modules used	0.015	<.001	2.70	0.009	<.001	1.62

[a]UAR: user activity ratio.

[b]ME: marginal effect reflecting the change in the outcome as the continuous variable changes by one unit (from 0 to 1 for categorical variables).

Table 5. Beta regressions estimating the effect of user characteristics on user engagement (user activity ratio) for (partly) automated modules.

User characteristics	App module								
	Exercise (n=1579; average UAR[a]=67 days)			BG[b] readings (n=5015; average UAR=13 days)			CGM[c] (n=559; average UAR=99 days)		
	ME[d]	*P* value	Days	ME	*P* value	Days	ME	*P* value	Days
Sociodemographic									
Gender (reference=male)									
Female	0.059	.003	10.62	−0.019	<.001	−3.42	−0.082	.02	−14.76
Unknown	0.006	.80	10.80	−0.024	<.001	−4.32	−0.058	.15	−10.44
Age (reference=30-40 years)									
Below 30 years	−0.067	.009	−12.06	0.002	.54	0.36	−0.026	.58	−4.68
40-50 years	−0.008	.68	−1.44	0.007	.004	1.26	0.035	.39	6.30
50-60 years	0.004	.85	0.72	0.015	<.001	2.70	0.022	.57	3.96
60-70 years	0.040	.09	7.20	0.018	<.001	3.24	0.008	.87	1.44
Above 70 years	0.032	.39	5.76	0.033	<.001	5.94	0.090	.14	16.20
Unknown	−0.078	.37	−14.04	0.002	.77	0.36	0.068	.15	12.24
BMI (reference=normal)									
Underweight (BMI<18.5)	0.180	.64	32.4	−0.021	.30	−3.78	0.081	.54	14.58
Overweight (25≤BMI<30)	0.019	.63	3.42	−0.002	.72	−0.36	0.039	.40	7.02
Obese (30≤BMI)	0.003	.93	0.54	−0.004	.53	−0.72	0.007	.88	1.26
Unknown	−0.002	.94	−0.36	−0.008	.20	−1.44	0.050	.29	9.00
Diagnosis									
Type (reference=type 1 diabetes)									
Type 2 diabetes	−0.006	.80	−1.08	0.003	.28	0.54	−0.024	.55	−4.32
Prediabetes	−0.033	.36	−5.94	−0.003	.61	−0.54	−0.216	.12	−38.88
Other	−0.038	.31	−6.84	0.006	.27	1.08	−0.071	.63	−12.78
Unknown	0.022	.55	3.96	0.046	<.001	8.28	0.058	.62	10.44
Recently diagnosed (reference=no)									
Yes	−0.021	.22	−3.78	0.005	.006	0.90	0.059	.045	11.34
Unknown	0.013	.75	2.34	0.005	.42	0.90	0.204	.17	36.72
Disease status									
Treatment (reference=noninsulin)									
Insulin treatment	−0.005	.76	−0.90	−0.006	.002	−1.08	0.100	.13	18.00
Unknown	−0.035	.27	−6.30	0.000	.98	0.00	−0.167	.14	−30.06
BG in control (reference=no)									
Yes	−0.014	.49	−2.52	0.001	.43	0.18	−0.029	.61	−5.22
Unknown	−0.045	.02	−8.10	−0.031	<.001	−5.58	−0.036	.31	−6.48
Usage									
Number of modules used	0.029	<.001	5.22	0.015	<.001	2.70	−0.002	.88	−0.36

[a]UAR: user activity ratio.

[b]BG: blood glucose.

[c]CGM: continuous glucose monitoring.

[d]ME: marginal effect reflecting the change in the outcome as the continuous variable changes by one unit (from 0 to 1 for categorical variables).

Discussion

Principal Findings

The aims of this study were to investigate how intensively users engage with mHealth for diabetes and to identify patient characteristics that are associated with user engagement. Our study revealed that 42.79% (12,685/29,643) of individuals who gave consent initiated use, thus reflecting effective transition from download to use. This could reflect a superficial exploration of several apps or a general insecurity on whether and how to best integrate mHealth apps into disease self-management. The superficial exploration is likely reinforced by the large availability of different apps in the diabetes field. Furthermore, the use may be impacted by the price of the app (the app used in this study is free of charge), and if the app is selected in collaboration with an HCP [27,28]. A survey among patients with diabetes in China revealed that only 19% of the apps were recommended by HCPs, whereas most users selected their app randomly or as recommended by other patients [28]. Finally, users may be reluctant to share information due to privacy concerns. It was found that a large proportion of users did not fill out optional fields in the self-reported user profiles or entered implausible values.

In contrast to the National Diabetes Statistics Report 2020, according to which men had a higher prevalence of diagnosed diabetes, 56.86% (2075/3649) of users reported to be female [29]. This could either indicate that women are more likely to use the app or that women are more willing to share information on their gender. A total of 13.48% (1220/9051) of the users were diagnosed with T1D. Estimates from the ADA suggest that only about 5% of people with diabetes are diagnosed with T1D [30]. Explanations for the high share of users with T1D in our study may either be their higher need for constant self-monitoring (and thus demand for the app) or their usually younger age of disease onset.

For the users who initiated use, the activity delay was much higher on average for the CGM module than for the other modules. This could be explained by the increasing availability of CGM technology during the period when data for this study were collected. Therefore, some users may have gained access to a CGM device months after they consented to the app. In addition, effort and difficulty of use have been identified as the main reasons for dissatisfaction among mHealth users in a previous study [31]. Hence, the activity delay may reflect the burden of initiating the use of the CGM module, for example, due to the necessity to connect the app with an external medical device.

The average 180 days UAR varied between 0.05 and 0.55, depending on the module of the app. Comparing these numbers with existing studies is difficult because engagement metrics and reporting remain heterogeneous in the literature. However, a discrepancy between initial uptake and long-term use was found among the modules: we observed either a fast and high uptake with a lower long-term use or a lower and later uptake and use that is more continuous. Low or decreasing engagement has also been reported in previous studies among patients with diabetes (it should be noted that high churn rates have also been observed for apps in other categories and are not health or disease-specific) [13,17,19,20,32-37]. A recent study using an in-app embedded questionnaire revealed that the most satisfying user experiences took place within the first week of engagement and were related to visual elements and the feasibility of health monitoring [31]. Thus, it is a well-known challenge to keep the user engaged over time.

One reason for the dropouts in this study may be suboptimal service matching, that is, the users engage with the app but experience that the app may not fit their exact needs [38]. The C4C app offers modules targeting different aspects of diabetes treatment. The app could thus demonstrate to patients how dosing along with exercise and diet affects their diabetes. However, most users only take advantage of one specific module, which may suggest that user needs are not the same and that most users may not need the full set of functionalities. Moreover, the UAR was highest for the CGM module, that is, a new technology with potentially fewer app suppliers (ie, alternative products) on the market.

Another cause for low long-term use may be that the app is not used optimally and may therefore result in a suboptimal user experience, for example, due to missing awareness about the importance of (true) entries for effective disease management. Therefore, a rewarding experience may be crucial to keep the users who are less aware of the importance of disease monitoring engaged in the long term [31]. In contrast, it may as well be possible that a patient is reassured about his disease management after having used the app for a while and does not *need* it anymore.

In this study, the discrepancy between high uptake and fast discontinuation was particularly noticeable in the BG module, in the medication module, and in the food module. In this context, it should be emphasized that the mode of data collection varies between modules. The exercise and CGM modules automatically collect data once the app is connected to an external fitness app or the CGM device. Therefore, exploiting the module does not require any further activity from the users than opening the app. The passive data collection may thus explain the high UAR for these modules. In contrast, users must actively make entries in the medication, food, and BG modules (the number of tabs can vary depending on the exact way they choose), that is, modules with lower engagement. Accordingly, the median number of seconds to perform an entry was lowest in the CGM module (0 seconds) and in the exercise module (0 seconds, if from a connected app), followed by the BG module (58 seconds), the medication module (117 seconds), and the food module (156 seconds; Multimedia Appendix 1). Therefore, the mode of data collection may be an important driver for engagement with the app. This finding is in accordance with the existing literature based on surveys. The studies revealed that the data entry burden is the main reason for discontinuation of use, and the authors suggest that enabling features with automatically transmitted data would increase the users' compliance [27,28]. However, it should be pointed out that it

remains unclear how actively users ultimately interact with the app when data have been transmitted automatically. As the app needs to be opened to sync, the engagement metrics can (at least) serve as a proxy for user engagement. Furthermore, besides engagement, the metrics inform about the amount of data collected and reflect the availability of rich health information. Previous studies showed that both patients and diabetologists thought that diaries and doctor-patient communication based on collected data were the most important features of apps for diabetes [28,39]. On the other hand, we found that the medication, food, and BG modules had a higher proportion of users than the other modules, potentially because the manual data entry mode makes it possible for any person with diabetes to register data (for the BG module, a BG meter is necessary, but these are very common for people with diabetes). In contrast, the CGM and exercise modules are meant for synchronization with devices or apps, which the user may not have or which pose an additional burden and complexity to initiate use.

Our regression results indicate that the optimal approach to increase user engagement may be patient specific and depend on their individual characteristics. In line with the findings of Rahman et al [18], we concluded that although most users of the app were female, male users were significantly more likely to be engaged in the app. Only the exercise module was used significantly more by female users. This result could be explained by a higher body consciousness, which is in accordance with existing literature revealing that women reported higher exercise levels than men [40]. Another explanation may be that exercise data are often passively collected and would require active disconnection from the external app. Furthermore, we showed that older users engaged significantly more with the app. Compared with previous studies, this finding is rather novel. On the basis of the interviews with patients aged 50 years or older, Scheibe et al [41] revealed a lack of acceptance and a low use of diabetes apps among the elderly, mainly due to a lack of additional benefits and ease of operation. Similarly, using survey data, Zhang et al [28] showed that the use of diabetes apps decreased with patient age. One explanation for our finding may be that if older individuals overcome a potential lack of knowledge in app use, they may be motivated to apply their newly obtained skills. In contrast, younger mobile device owners might consider their phone as a spare time activity and have a lower burden downloading and trying different apps [42]. Moreover, considering that diabetes is a chronic condition, the older may be more severely ill, and therefore, they require more (self-) monitoring compared with younger patients in earlier stages of the disease.

In addition, our analysis reveals that there is a tendency that individuals with a more recent first diabetes diagnosis used the app more intensively. So far, the existing literature could not find a correlation between disease duration and app use [28]. However, it is possible that new patients have considered their mobile phone as a helpful tool when establishing their initial individual self-management strategy. The great potential of patient support apps is increasingly emphasized by official institutions, such as the ADA or the WHO [11,43]. In contrast, long-term patients may not consider adjusting their predeveloped disease management habits. A usability study with older adults showed that participants were already satisfied with their current management system and that they would need to find a reason why apps are superior to their current medication management system [44]. However, it should be noted that our observation is only based on an approximate measure for the time of diagnosis, and more research is needed to confirm this finding.

The total number of actively used modules was positively associated with the interaction in a single module. This could indicate synergy effects among the modules, if the app was integrated in managing more aspects of a user's disease or could reflect a higher discipline of users who take advantage of multiple modules.

Limitations

Our study has several limitations. First, missing data is a common problem when analyzing self-reported real-world data, ranging from 9.88% (894/9051) of users who did not share information about their diabetes type to 63.31% (5730/9051) of users who did not share information about their height or weight. We addressed this by categorizing missing data as unknown. In the context of this study, the amount of missing data can be viewed as an important result in itself. It is noteworthy that many unknown categories are highly significant. This may indicate that there are factors associated with user engagement that could not be observed in the given data set. Second, in a broader sense, data limitations constrained our ability to investigate other important factors that may influence user engagement, such as education or patient-doctor interactions. Wherever possible, we generated proxies substituting unobserved variables of interest. Finally, the data set did not cover information about the log-in times. Instead, our study was solely based on data inputs. Although we emphasized the relevance of automated data collection in our discussion, it remains unclear how actively engaged the users actually are, because collecting data is not the same as using data. In this context, it should be the ultimate goal of future research to assess how engagement with mHealth affects health outcomes and which module is most important from the health perspective.

Conclusions

There is no consensus on how to assess user engagement with mHealth solutions, but it is needed. This paper proposes a set of theoretically founded user engagement metrics that were used to assess user engagement in this study. After providing consent, 42.79% (12,685/29,643) of the users initiated use. Most users took advantage of one specific module of the app, indicating that the needs of patients with diabetes are highly heterogeneous. User engagement and amount of collected data were higher for automated modules, although initial uptake remained lower for these modules. Therefore, to increase the use of apps, providers of mHealth should consider the mode of data gathering and content design and take into account the privacy concerns of the users at the same time. Users' engagement was determined by various patient characteristics. Although most users reported to be female, male users engaged significantly more with the app. Older people and users who were recently diagnosed tended

to use the app more actively. This indicates that particular patient groups should be specifically targeted or assisted when integrating apps into the self-management of their disease.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement No. 721402. The authors gratefully thank Jessica Baldridge (Glooko) and Katarina LeFever (Novo Nordisk) for providing background information about the app. The authors gratefully thank Professor Werner Brouwer (Erasmus University Rotterdam) and Professor Aleksandra Torbica (Bocconi University) for their suggestions and comments on the manuscript.

Authors' Contributions

AB, TS, and MJ developed the idea of the study. AB, MS, MJ, and TS developed the design of the study. AB and MS contributed to data management and descriptive statistics. AB performed the regression analysis. AB drafted the manuscript. All authors contributed to the interpretation of the results and read and approved the final version of the paper.

Conflicts of Interest

Cornerstones4Care is a registered trademark of Novo Nordisk A/S. MJ and MS are full-time employees of Novo Nordisk A/S, and AB is associated with Novo Nordisk A/S through a research collaboration as part of her PhD.

Multimedia Appendix 1
Users' burden to do an entry in each module.
[DOCX File , 13 KB - mhealth_v8i11e22212_app1.docx]

References

1. Alberti K, Zimmet P. 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]
2. American Diabetes Association. Standards of Medical Care in Diabetes—2020. Diabetes Care 43. Supplement 1. URL: https://care.diabetesjournals.org/content/43/Supplement_1 [accessed 2020-11-03]
3. Simpson SH, Lin M, Eurich DT. Medication adherence affects risk of new diabetes complications: a cohort study. Ann Pharmacother 2016 Sep;50(9):741-746. [doi: 10.1177/1060028016653609] [Medline: 27307411]
4. Bellei EA, Biduski D, Cechetti NP, de Marchi AC. Diabetes mellitus m-Health applications: a systematic review of features and fundamentals. Telemed J E Health 2018 Nov;24(11):839-852. [doi: 10.1089/tmj.2017.0230] [Medline: 29470105]
5. Kleinman NJ, Shah A, Shah S, Phatak S, Viswanathan V. Improved medication adherence and frequency of blood glucose self-testing using an m-Health platform versus usual care in a multisite randomized clinical trial among people with type 2 diabetes in India. Telemed J E Health 2017 Sep;23(9):733-740. [doi: 10.1089/tmj.2016.0265] [Medline: 28328396]
6. Karduck J, Chapman-Novakofski K. Results of the clinician apps survey, how clinicians working with patients with diabetes and obesity use mobile health apps. J Nutr Educ Behav 2018 Jan;50(1):62-9.e1. [doi: 10.1016/j.jneb.2017.06.004] [Medline: 29325664]
7. Wu Y, Yao X, Vespasiani G, Nicolucci A, Dong Y, Kwong J, et al. Mobile app-based interventions to support diabetes self-management: a systematic review of randomized controlled trials to identify functions associated with glycemic efficacy. J Med Internet Res Mhealth Uhealth 2017 Mar 14;5(3):e35 [FREE Full text] [doi: 10.2196/mhealth.6522] [Medline: 28292740]
8. Hou C, Carter B, Hewitt J, Francisa T, Mayor S. Do mobile phone applications improve glycemic control (HbA1c) in the self-management of diabetes? A systematic review, meta-analysis, and grade of 14 randomized trials. Diabetes Care 2016 Nov;39(11):2089-2095. [doi: 10.2337/dc16-0346] [Medline: 27926892]
9. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. Telemed J E Health 2012 Apr;18(3):175-184. [doi: 10.1089/tmj.2011.0119] [Medline: 22356525]
10. Blenner SR, Köllmer M, Rouse AJ, Daneshvar N, Williams C, Andrews LB. Privacy policies of android diabetes apps and sharing of health information. J Am Med Assoc 2016 Mar 8;315(10):1051-1052. [doi: 10.1001/jama.2015.19426] [Medline: 26954415]
11. Fleming GA, Petrie JR, Bergenstal RM, Holl RW, Peters AL, Heinemann L. Diabetes digital app technology: benefits, challenges, and recommendations. A consensus report by the european association for the study of diabetes (EASD) and the american diabetes association (ADA) diabetes technology working group. Diabetes Care 2020 Jan;43(1):250-260. [doi: 10.2337/dci19-0062] [Medline: 31806649]

12. Jimenez G, Lum E, Car J. Examining diabetes management apps recommended from a google search: content analysis. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11848 [FREE Full text] [doi: 10.2196/11848] [Medline: 30303485]
13. Ernsting C, Dombrowski SU, Oedekoven M, 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 5;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
14. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 2017 Apr 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
15. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of ehealth usage: insights on the digital divide from the health information national trends survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
16. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health app use among US mobile phone users: analysis of trends by chronic disease status. JMIR Mhealth Uhealth 2017 Dec 19;5(12):e197 [FREE Full text] [doi: 10.2196/mhealth.7832] [Medline: 29258981]
17. Serrano KJ, Coa KI, Yu M, Wolff-Hughes DL, Atienza AA. Characterizing user engagement with health app data: a data mining approach. Transl Behav Med 2017 Jun;7(2):277-285 [FREE Full text] [doi: 10.1007/s13142-017-0508-y] [Medline: 28616846]
18. Rahman QA, Janmohamed T, Pirbaglou M, Ritvo P, Heffernan JM, Clarke H, et al. Patterns of user engagement with the mobile app, manage my pain: results of a data mining investigation. JMIR Mhealth Uhealth 2017 Jul 12;5(7):e96. [doi: 10.2196/mhealth.7871] [Medline: 28701291]
19. Becker S, Kribben A, Meister S, Diamantidis CJ, Unger N, Mitchell A. User profiles of a smartphone application to support drug adherence: experiences from the iNephro project. PLoS One 2013;8(10):e78547 [FREE Full text] [doi: 10.1371/journal.pone.0078547] [Medline: 24194946]
20. Taki S, Lymer S, Russell CG, Campbell K, Laws R, Ong K, et al. Assessing user engagement of an mHealth intervention: development and implementation of the growing healthy app engagement index. JMIR Mhealth Uhealth 2017 Jun 29;5(6):e89 [FREE Full text] [doi: 10.2196/mhealth.7236] [Medline: 28663164]
21. Nelson LA, Coston TD, Cherrington AL, Osborn CY. Patterns of user engagement with mobile- and web-delivered self-care interventions for adults with T2DM: a review of the literature. Curr Diab Rep 2016 Jul;16(7):66 [FREE Full text] [doi: 10.1007/s11892-016-0755-1] [Medline: 27255269]
22. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
23. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J. Am. Soc. Inf. Sci 2008 Feb 28;59(6):938-955. [doi: 10.1002/asi.20801]
24. Peterson ET. Measuring the Immeasurable: Visitor Engagement. Vertical Studio. URL: http://www.verticalstudio.com/hs-fs/hub/74398/file-15425996-pdf/docs/%20web_analytics_demystified_and_nextstage_global_-_measuring_the_immeasurable_-_visitor_engagement.pdf [accessed 2020-10-20]
25. The Big Picture: Checking Your Blood Glucose. American Diabetes Association. URL: https://www.diabetes.org/diabetes/medication-management/blood-glucose-testing-and-control/checking-your-blood-glucose [accessed 2020-04-02]
26. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range. Diabetes Care 2019 Aug;42(8):1593-1603 [FREE Full text] [doi: 10.2337/dci19-0028] [Medline: 31177185]
27. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 4;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
28. Zhang Y, Li X, Luo S, Liu C, Xie Y, Guo J, et al. Use, perspectives, and attitudes regarding diabetes management mobile apps among diabetes patients and diabetologists in China: national web-based survey. JMIR Mhealth Uhealth 2019 Feb 8;7(2):e12658 [FREE Full text] [doi: 10.2196/12658] [Medline: 30735147]
29. National Diabetes Statistics Report, 2020. Estimates of Diabetes and Its Burden in the United States. Centers for Disease Control and Prevention. Atlanta, GA; 2020. URL: https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf [accessed 2020-04-02]
30. Statistics About Diabetes. American Diabetes Association. URL: https://www.diabetes.org/resources/statistics/statistics-about-diabetes [accessed 2020-05-26]
31. Biduski D, Bellei EA, Rodriguez JP, Zaina LA, de Marchi AC. Assessing long-term user experience on a mobile health application through an in-app embedded conversation-based questionnaire. Comput Hum Behav 2020 Mar;104:106169. [doi: 10.1016/j.chb.2019.106169]
32. 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]

33. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. J Med Internet Res 2011 Jan 25;13(1):e9 [FREE Full text] [doi: 10.2196/jmir.1391] [Medline: 21371992]
34. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
35. Park YR, Lee Y, Kim JY, Kim J, Kim HR, Kim Y, et al. Managing patient-generated health data through mobile personal health records: analysis of usage data. JMIR Mhealth Uhealth 2018 Apr 9;6(4):e89 [FREE Full text] [doi: 10.2196/mhealth.9620] [Medline: 29631989]
36. Percentage of Mobile Apps That Have Been Used Only Once From 2010 to 2019. Statistica. 2020. URL: https://www.statista.com/statistics/271628/percentage-of-apps-used-once-in-the-us/ [accessed 2020-07-06]
37. Chen A. New Data Shows Losing 80% of Mobile Users is Normal, and Why the Best Apps Do Better. Andrewchen. URL: https://andrewchen.co/new-data-shows-why-losing-80-of-your-mobile-users-is-normal-and-that-the-best-apps-do-much-better/ [accessed 2020-07-06]
38. Wang L, Wu T, Guo X, Zhang X, Li Y, Wang W. Exploring mHealth monitoring service acceptance from a service characteristics perspective. Electronic Comm Res Appl 2018 Jul;30:159-168. [doi: 10.1016/j.elerap.2018.06.003]
39. Zhang Y, Li X, Luo S, Liu C, Liu F, Zhou Z. Exploration of users' perspectives and needs and design of a type 1 diabetes management mobile app: mixed-methods study. JMIR Mhealth Uhealth 2018 Sep 21;6(9):e11400 [FREE Full text] [doi: 10.2196/11400] [Medline: 30249580]
40. Craft BB, Carroll HA, Lustyk MK. Gender differences in exercise habits and quality of life reports: assessing the moderating effects of reasons for exercise. Int J Lib Arts Soc Sci 2014 Jun;2(5):65-76 [FREE Full text] [Medline: 27668243]
41. Scheibe M, Reichelt J, Bellmann M, Kirch W. Acceptance factors of mobile apps for diabetes by patients aged 50 or older: a qualitative study. Med 2 0 2015 Mar 2;4(1):e1 [FREE Full text] [doi: 10.2196/med20.3912] [Medline: 25733033]
42. Haug S, Castro RP, Kwon M, Filler A, Kowatsch T, Schaub MP. Smartphone use and smartphone addiction among young people in Switzerland. J Behav Addict 2015 Dec;4(4):299-307 [FREE Full text] [doi: 10.1556/2006.4.2015.037] [Medline: 26690625]
43. Mobile Health for Diabetes Prevention and Management (mDiabetes). World Health Organization. 2016. URL: https://www.who.int/ncds/prevention/be-healthy-be-mobile/hanbook-mDiabetes/en/ [accessed 2020-09-16]
44. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014 Mar 14;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]

Abbreviations

ADA: American Diabetes Association
BG: blood glucose
C4C: Cornerstones4Care
CGM: continuous glucose monitoring
HCP: health care professional
mHealth: mobile health
T1D: type 1 diabetes
T2D: type 2 diabetes
TIR: time in range
UAR: user activity ratio
WHO: World Health Organization

Edited by G Eysenbach; submitted 06.07.20; peer-reviewed by L Nelson, E Bellei; comments to author 15.08.20; revised version received 24.09.20; accepted 24.09.20; published 06.11.20.

Please cite as:
Böhm AK, Jensen ML, Sørensen MR, Stargardt T
Real-World Evidence of User Engagement With Mobile Health for Diabetes Management: Longitudinal Observational Study
JMIR Mhealth Uhealth 2020;8(11):e22212
URL: http://mhealth.jmir.org/2020/11/e22212/
doi:10.2196/22212
PMID:32975198

Original Paper

Use of the patientMpower App With Home-Based Spirometry to Monitor the Symptoms and Impact of Fibrotic Lung Conditions: Longitudinal Observational Study

Colin Edwards[1], PhD; Eamonn Costello[1], BEng; Nicola Cassidy[2]; Bill Vick[3]; Anne-Marie Russell[4,5], PhD, MSc, ATSF

[1]patientMpower Ltd, Dublin, Ireland

[2]Irish Lung Fibrosis Association, Dublin, Ireland

[3]PF Warriors, Plano, TX, United States

[4]Imperial College Healthcare NHS Trust, London, United Kingdom

[5]University of Exeter College of Medicine and Health, Exeter, United Kingdom

Corresponding Author:
Colin Edwards, PhD
patientMpower Ltd
21 Denzille Lane
Saint Peter's
Dublin, D02 EY19
Ireland
Phone: 353 872599131
Email: colin@patientmpower.com

Abstract

Background: Daily home-based spirometry in idiopathic pulmonary fibrosis (IPF) has been shown to be feasible and clinically informative. The patientMpower app facilitates home-based spirometry along with home-based monitoring of IPF-related symptoms. The patientMpower app can be downloaded to the user's mobile phone or tablet device, enabling the recording of objective and subjective data.

Objective: The aim of this paper is to report on the 1-year experience of using patientMpower with home-based spirometry by 36 participants with self-reported pulmonary fibrosis (PF) treated with usual care.

Methods: Self-selecting participants enrolled in this community-based participatory research program through a patient advocacy group in their country: Irish Lung Fibrosis Association in Ireland and PF Warriors in the United States. Disease severity was comparable with a baseline mean predicted forced vital capacity (FVC) of 64% and 62% in the Irish and US participants, respectively. Both groups of participants were allocated to identical, in-country, open-label, single-group observational studies and were provided with a Bluetooth-active Spirobank Smart spirometer integrated directly with patientMpower. Data collected via patientMpower included seated FVC (daily), breathlessness grade (modified Medical Research Council scale score), step count, medication adherence, and symptoms and impact of IPF on daily life, which were measured by a patient-reported outcome measure (PROM) scale that was specifically developed for IPF. Longitudinal patient-reported data on oximetry and oxygen consumption were also collected.

Results: A large majority of the 36 participants reported that their experience using patientMpower was positive, and they wanted to continue its use after the initial 6-week observation. Out of 36 participants, 21 (58%) recorded home-based spirometry without prompting for ≥180 days, and 9 (25%) participants continued with recording home-based spirometry for ≥360 days.

Conclusions: The patientMpower app with associated Bluetooth-connected devices (eg, spirometer and pulse oximeter) offers an acceptable and accessible approach to collecting patient-reported objective and subjective data in fibrotic lung conditions.

(JMIR Mhealth Uhealth 2020;8(11):e16158) doi:10.2196/16158

KEYWORDS

idiopathic pulmonary fibrosis; pulmonary fibrosis; eHealth; mHealth; patient-reported outcome measure; spirometry; home spirometry; patient experience; digital health; patient advocacy

Introduction

Pulmonary fibrosis (PF) is an irreversible lung disease that leads to progressive breathlessness and deterioration of pulmonary function; it is associated with significant limitation of physical activities and worsening quality of life [1]. The majority of PF cases are idiopathic, and there is no cure. The advent of disease-modifying/antifibrotic pharmacotherapies in the last decade has slowed the progression of PF, with registry data reporting an increase in median survival from 3 to 4.5 years. The progression of PF is highly variable. Acute exacerbations resulting in further impairment of lung function are a major cause of morbidity and mortality in idiopathic pulmonary fibrosis (IPF), reducing median survival to <21 months where hospitalization is needed [2].

Adjunctive symptom-based management is integral to disease management given the high symptom burden [3]. Objective assessments of lung function and symptom experience (including breathlessness and fatigue) are important in the monitoring of patients with IPF to inform both their self-management and hospital-based care plan, particularly initiation and adjustment of oxygen therapy, pulmonary rehabilitation, palliative care, and psychological therapies [4]. Traditional monitoring of IPF has typically involved spirometric measurement of forced vital capacity (FVC) along with the assessment of IPF-related symptoms at 6- and 12-month intervals in outpatient clinical settings.

Digital platforms which directly record patients' symptom experiences and impacts of those symptoms on patients' daily lives alongside objective clinical measurements (eg, blood pressure) have been developed for chronic medical conditions to support patients in managing their health. There was a lack of digital platforms created specifically for patients with IPF, and there were few disease-specific instruments to assess health-related quality of life with IPF. The IPF patient-reported outcome measure (IPF-PROM) captures symptom experience and impact of IPF on patients' daily lives [5]. The collection of serial IPF-PROM data provides valuable information on longitudinal trends at individual and population levels. Furthermore, home-based monitoring of FVC by patients using small handheld spirometers is feasible, clinically informative, and more accurate in estimating lung function in patients with IPF than intermittent spirometry performed at clinic visits, and it may be of value as a primary endpoint in short proof-of-concept studies in IPF [6-8].

The patientMpower app was developed to facilitate home-based spirometry and symptom monitoring in PF. Users download patientMpower to their mobile phones or tablet devices, enabling them to record both objective data (eg, FVC) and subjective data (eg, impact on quality of life) simultaneously. There are few published data on longitudinal trends in patient-measured FVC and patient-reported outcome measures (PROMs) in IPF. Patient-reported data may predict important health outcomes (eg, exacerbations). We have previously reported on 6-week feasibility studies of patientMpower in patients with PF [9,10] These studies were codesigned in collaboration with patient advocacy groups (in Europe and the United States). Here, we present the 1-year follow-up of these studies designed to test the acceptability and utility of patientMpower.

Methods

We conducted 2 parallel studies of the same design and methodology in 2 distinct populations (in the United States and Ireland). The studies were codesigned and conducted in collaboration with 2 advocacy groups for patients with PF: PF Warriors, Texas, United States, and Irish Lung Fibrosis Association (ILFA), Dublin, Ireland. Our approach was not prescriptive, as we aimed to observe the real-world experience of the integration of patientMpower into the daily lives of people with fibrotic lung conditions.

The patient advocacy groups collaborated with us from the outset of this study, informing, reviewing, and approving the study concept and design and leading on the patient information documents prior to the initiation of the study. The study protocol evolved iteratively, informed by the stakeholder and participant discussions and by peer review by health care professionals. Our approach was not that of conventional science and was rather founded in an implementation science approach, working "with" participants instead of "on" them. Participatory action research seeks to empower participants to tailor an intervention to suit their own contexts [11]. We were keen to ensure that patientMpower fitted into the context of an individual's life. When using such approaches, the area between engagement/consultation and research is gray. Given that this work was led by patient advocacy groups and was focused on increasing experiential knowledge and obtaining results within a timeframe to determine how to scale up the use and sustainability of patientMpower, we were informally advised that approval from an institutional review board was not required. Nonetheless, in line with principles of good clinical practice, all participants provided written informed consent prior to accessing patientMpower and participating in any study-specific procedures. Entry criteria included a diagnosis of PF, ownership of a smartphone or tablet device, an email address, access to the internet at home, and willingness to provide written informed consent.

Each study was an open-label, single group, observational study in patients with PF treated with usual care. After obtaining consent, participants were invited to download the patientMpower app and use it daily for ≥6 weeks (the active observational study period). All participants were provided with a Spirobank Smart spirometer (Medical International Research), which integrates directly via Bluetooth with patientMpower to capture spirometry data. FVC data were displayed to each participant on their mobile device. Video-based training on the correct use of patientMpower and home-based spirometry was provided on a dedicated YouTube channel [12]. Participants were asked to record the following data via the app: FVC (one forced expiratory maneuver per day, seated), breathlessness (using the modified Medical Research Council [mMRC] score, once a day), step count, medication adherence, symptoms, and impact of IPF on daily life (using PROM scores) [5]. The IPF-PROM consisted of an overall quality of life question and 12 questions covering 4 domains: physical impacts of

breathlessness, psychological impacts of breathlessness, psychological well-being, and fatigue. There were no prompts to record spirometry or breathlessness; however, participants were prompted to record a PROM once per week for the first 6 weeks.

Participants in the United States could record pulse oximetry on patientMpower using a Nonin 3230 Bluetooth-linked pulse oximeter (Nonin Medical Inc) and oxygen flow rates. The protocol did not specify the frequency or conditions for recording pulse oximetry, and the pulse oximeter was not supplied by the study sponsor.

At the end of the initial 6-week observation, participants were asked to complete a questionnaire to assess the utility and acceptability of patientMpower (Multimedia Appendix 1). Thereafter, participants were free to use patientMpower and the spirometer if they chose.

The sample sizes for each study were determined by estimating the number of participants that could be recruited by the advocacy groups over a 2-month period. We expected 50 participants to be enrolled by the PF Warriors and 30 participants from the ILFA.

The results are tabulated and discussed in the following section. Anonymized demographic data are displayed for all participants who gave informed consent. All other displayed data are for participants who used patientMpower and recorded spirometry at least once.

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

Results

Study Population

In the United States, 27 participants enrolled in the study; 24 used patientMpower and recorded spirometry at least once. In Ireland, 13 participants enrolled in the study, all used patientMpower, and 12 recorded spirometry at least once. Baseline demographics were comparable in both groups (Table 1). Most US study participants (25/27, 93%) reported a confirmed diagnosis of PF; this information was not reported by Irish study participants. Most US participants (23/27, 85%) and some Irish participants (4/13, 31%) reported use of antifibrotic therapy at baseline.

Table 1. Baseline demographic data.

Characteristics	US study participants (N=27)	Ireland study participants (N=13)
Gender, n (%)		
Male	12 (44)	7 (54)
Female	13 (48)	6 (46)
Not stated	2 (7)	0 (0)
Ethnicity, n (%)		
White	24 (89)	13 (100)
Other	1 (4)	0 (0)
Not stated	2 (7)	0 (0)
Age (years), mean (range)	62 (31-79)	66 (37-81)
FVC[a,b] (L), mean (range)	2.28[c] (0.6-4.72)	2.49[d] (0.7- 3.9)
FVC[b] (% predicted), mean (range)	62 (36-108)	64 (33-91)
Diagnosis confirmed by clinical expert[e], n (%)		
Yes	25 (93)	0 (0)
No	0 (0)	0 (0)
Not stated	2 (7)	13 (100)
On antifibrotic therapy[e], n (%)		
Yes	23 (85)	4 (31)
No	0 (0)	0 (0)
Not stated	4 (15)	9 (61)

[a]FVC: forced vital capacity.

[b]Home-based spirometry data.

[c]Mean of first 7 days in 23 US participants. A total of 24 US participants provided home-based spirometry data; however, 1 participant did not provide spirometry data in the first 7 days.

[d]Mean of first 7 days in 12 Irish participants who provided home-based spirometry data.

[e]Self-reported data.

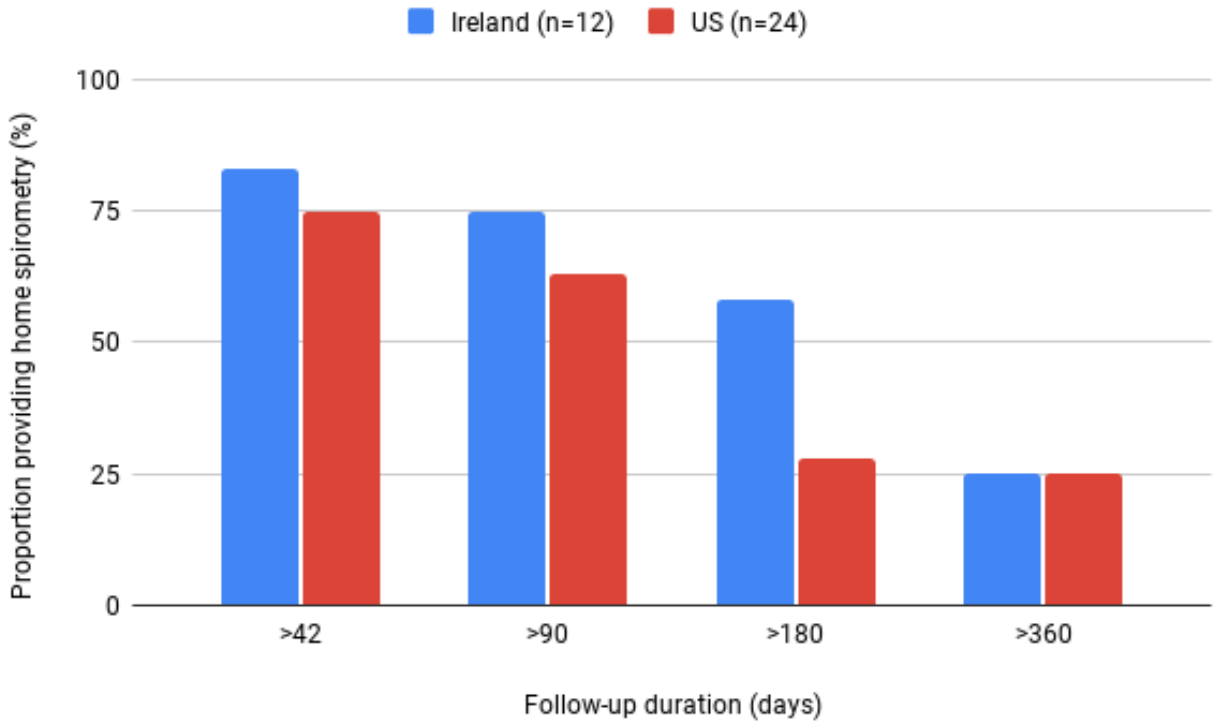
User-Reported Experience at 6 Weeks

At 6 weeks, 75% (18/24) of the US participants and 92% (11/12) of the Irish participants completed a questionnaire. Analysis of the 29 completed responses indicated that a large majority of respondents liked using patientMpower (28/29, 96%) and found it easy to use (26/29, 90%). A majority reported that patientMpower helped them take the correct dose of medication at the prescribed time and achieve personal exercise goals (19/29, 66%). Most reported that the effect of using patientMpower on their well-being and daily life was positive (27/29, 93%) and that it was useful to be able to record the impact of lung fibrosis on their well-being and daily life (25/29, 86%). Most participants stated they would recommend patientMpower to others and wanted to continue using it after the initial 6 weeks (27/29, 93%).

Frequency of Home-Based Spirometry Recording

A total of 24 US participants and 12 Irish participants recorded home-based spirometry, and most (28/36, 78%) continued home-based spirometry after the initial 6-week observation (Figure 1). Out of 36 participants, 21 (58%) recorded home-based spirometry at least once after 180 days, and 9 (25%) recorded it at least once after 360 days. There was variation in the frequency and duration of recording of home-based spirometry, with some participants recording spirometry daily and others recording intermittently but over a long duration of follow-up (data not shown). A total of 39% (14/36) of participants recorded home-based spirometry at least once a week on ≥50% of weeks over 1 year.

Figure 1. The proportion of participants from the Ireland and US studies recording home-based spirometry at least once against the duration of follow-up. Proportion of participants providing home-based spirometry data at each follow-up duration calculated relative to total numbers of participants who provided any spirometry data in that population. US: United States.



Recording Health Outcomes

A total of 22 US and 13 Irish participants completed at least 1 IPF-PROM. There was a trend toward a greater number of IPF-PROMs reported by US participants compared with Irish participants (mean 12/person vs 7/person, respectively). More US participants (13/22, 59%) provided IPF-PROMs after 180 days compared to Irish participants (3/13, 23%).

Breathlessness data (using mMRC score) were provided by 15 US participants and 11 Irish participants. Most participants reported breathlessness at baseline (median mMRC score for Irish participants=1; US participants=3) with only a few (2/11 in the Irish study and 2/15 in the US study) reporting an mMRC score of 0 (ie, dyspnea only with strenuous exercise). A greater number of breathlessness data reports were recorded by US participants compared to Irish participants (mean 28/person in the US study vs 12/person in the Ireland study). A similar proportion of participants in the Irish and US study populations recorded a breathlessness score at least once after 180 days (6/15 US participants, 40%, vs 5/11 Irish participants, 45%). There were no reports of exacerbation of IPF or death in the study populations.

Pulse Oximetry

Pulse oximetry data were not available for the Irish population, as Bluetooth integration of the pulse oximeter with patientMpower had not been implemented at that time. In the US study, 20 participants recorded pulse oximetry. The mean number of recordings was 58/person over a duration ranging

from 1 to 409 days. A total of 10 participants recorded oximetry for ≥180 days.

Figure 2 shows an example of multiple data recorded by a single US participant over approximately 400 days. This participant's self-recorded baseline FVC was 65% predicted, and they reported antifibrotic therapy (nintedanib 150 mg twice daily). This participant reported breathlessness (mMRC score 1 or 2) with a regular use of oxygen (flow rate 3 or 4 L/min) throughout. They described their overall quality of life in the PROM as "good" throughout the follow-up.

Figure 2. Patient-recorded spirometry, oximetry, and PROM in a single participant over 1 year in 4 domains (A) psychological impact of breathlessness, (B) psychological well-being, (C) physical impact of breathlessness, and (D) fatigue. FVC: forced vital capacity; PROM: patient-reported outcome measure.

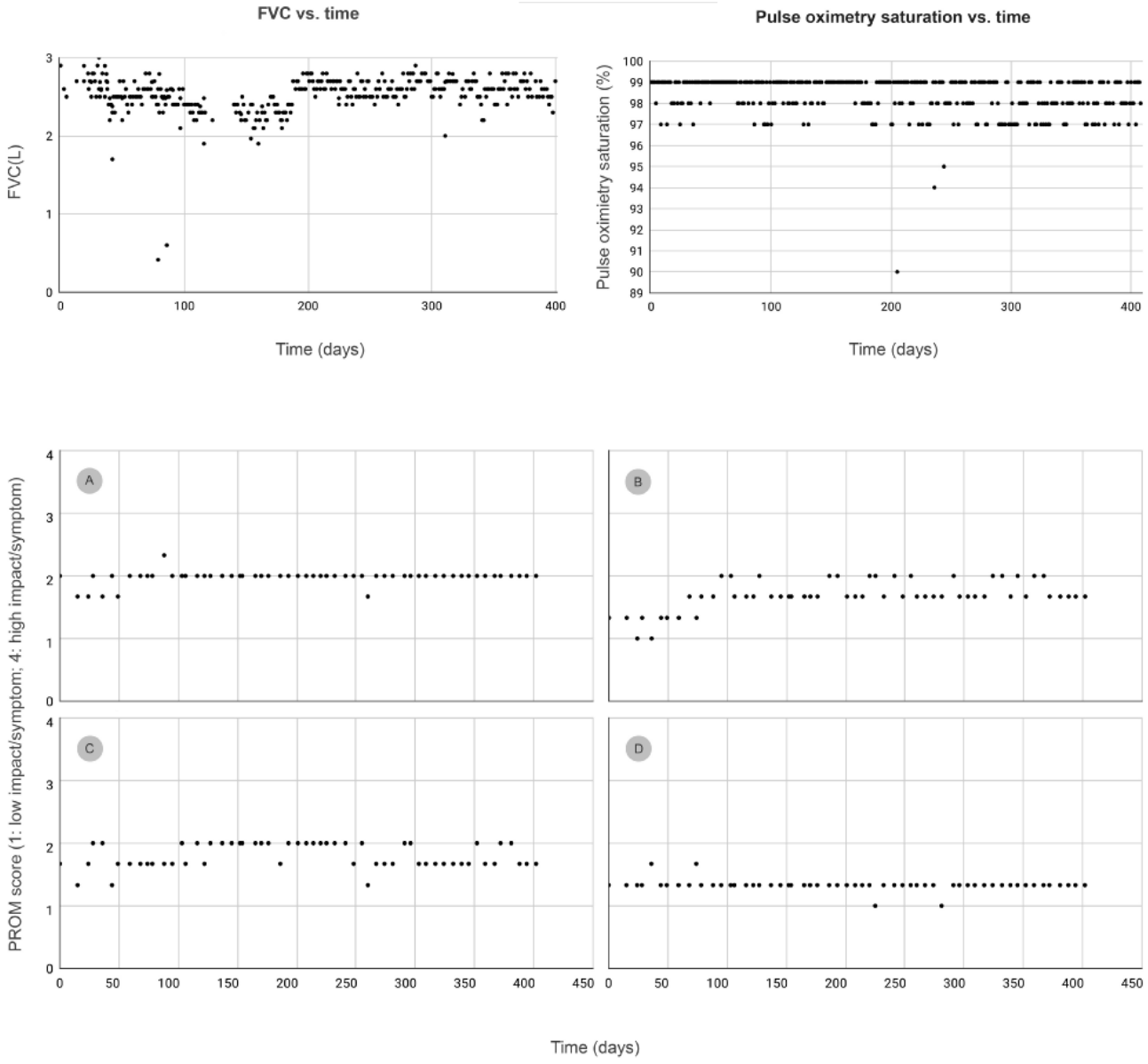


Figure 2 displays all patient-recorded values of FVC and pulse oximetry with no exclusion of outliers. Data on breathlessness (using mMRC scores) and oxygen flow rates were also collected (however, these are not shown in Figure 2). There were 4 grades of severity of a symptom or an impact of a symptom on quality of life for each question in the PROM. A score of 1 indicates lowest severity or impact (ie, the patient experienced the symptom/impact "none of the time" in the previous week). A score of 2 implies that the symptom/impact was experienced "some of the time," and a score of 3 implies that the symptom/impact was experienced "most of the time." A score of 4 indicates highest severity or impact (ie, the symptom/impact was experienced "all of the time" in the previous week). The scores for the 3 questions were averaged for each domain (psychological impact of breathlessness in panel A, psychological well-being in panel B, physical impact of breathlessness in panel C, and fatigue in panel D).

Discussion

Patients with chronic conditions (including PF) may be interested in using electronic tools to help them monitor and manage their health. The results we observed are consistent with those reported in a study in the Netherlands [13] in which a sample of patients with IPF were asked if they would welcome

a web-based platform to track their health; 67 (82%) responded "yes." The Dutch web-based platform was evaluated in a study [14] with 27 patients with IPF who recorded medication use and PROM at baseline and after 14 days. A large majority of these patients were "very satisfied" with their experience of this platform. This research team [14] went on to develop a home-based monitoring program using the MIR Spirobank Smart spirometer (the same device used in our studies); however, the FVC data were displayed on the patient's personal computer rather than on their mobile device. The FVC data were evaluated in 10 patients with IPF who recorded daily home-based spirometry and symptoms for 4 weeks. Adherence over 4 weeks was very high (99%), with good correlation between home-based and clinic spirometry records. All patients reported that home-based spirometry was useful for them; we replicate these findings here in larger numbers.

Patient-reported spirometry in patients with IPF has been evaluated and found to be feasible in other studies and, in general, it showed good correlation with clinic-measured spirometry [6,7]. In one study [6], correlation was observed between clinic spirometry at 3, 6, and 12 months, and the rate of decline of patient-recorded FVC was predictive of clinical outcome.

In another study, weekly home-based spirometry was evaluated in 25 patients with IPF using the best of 3 forced expiratory maneuvers with data captured directly by the spirometer (rather than in a paper diary) [7]. Mean adherence to weekly spirometry was 90.5% over the 24-week observation period, decreasing thereafter. The study [7] also reported good correlation between patient-measured and clinic spirometry data. The authors hypothesized that use of home-based patient-recorded spirometry data could improve the analytical efficiency of clinical trials and reduce the sample size required (vs clinic spirometry). This has proven to be of particular importance in the era of COVID-19.

The patientMpower app with integrated home-based spirometry has been used in clinical studies in patients with interstitial lung disease. In the first pilot-scale clinical study (n=7) of patientMpower for patients with IPF (recruited through a specialist interstitial lung disease clinic), there was a good correlation between home-based and clinical FVC recordings at 8 weeks [15]. At 8 months after the conclusion of that study, 3 of the 6 participants were still using patientMpower regularly. Other clinical studies in populations with interstitial lung disease reported good adherence to use of the patientMpower app with home-based spirometry [16,17].

There are several limitations to our study. The sample size was chosen arbitrarily (based on estimates of expected recruitment), and actual recruitment was lower than anticipated. Participants self-reported their diagnosis of PF, medication history, and other demographic data without clinical confirmation of these data. The participants were recruited through patient advocacy groups and were likely to represent highly engaged patients with a strong interest in monitoring and managing their health. The level of engagement demonstrated by the participants may not be typical of the general population of patients with PF. The questionnaire used to assess participants' opinions was developed de novo by the study sponsor and has not been tested in other populations.

A possible future research use of patientMpower would be the assessment of long-term patient-recorded spirometry and quality of life to better understand the correlation between lung function and subjective outcomes in chronic respiratory conditions. Inclusion of occasional formal clinic assessment of lung function and exercise capability (eg, 6-minute walking distance) would enhance the interpretation of study data (although in-clinic assessments are now much more difficult to access during the COVID-19 pandemic). The patientMpower app can be adapted so that clinical centers can view their patients' recorded data in real time to create a remote patient-monitoring approach. The patientMpower app with remote monitoring of pulse oximetry has already been implemented in lung transplantation centers and in a national program to support early discharge of patients diagnosed with COVID-19 from in-hospital care [18].

The age profiles of the patients in all of these studies were typical of IPF, implying that age per se is not a barrier to the use of home-based spirometry or electronic platforms to track health outcomes. To the best of our knowledge, our studies [9,10,15,17] are the first to report use of a specific mobile app by patients with IPF to capture and view their health data directly within a single platform. Patients with IPF appear to be motivated to record home-based spirometry regularly, even when not prompted.

Recruiting volunteers through patient support groups to participate in observational studies is feasible. Our work has shown that electronic informed consent and remote installation of health care apps (with associated sensor devices) can be implemented in observational studies for some patient populations with PF. This approach may be useful to capture patient-reported long-term trends in FVC, quality of life, and health outcomes in patients with PF and generate richer real-world evidence. This is illustrated by Figure 2 where the participant recorded objective data (FVC and oximetry) and subjective data (breathlessness and impact of IPF on quality of life) simultaneously at multiple time points over 1 year. The patientMpower app can also capture data on step count (from the user's smartphone or a connected FitBit device), enabling future analysis of activity levels in addition to the reported outcomes. To the best of our knowledge, presently, patientMpower is the only mobile app designed specifically for patients with PF that enables capture of all PF-relevant health outcomes within a single app.

Long-term use of the patientMpower app with integrated Spirobank Smart spirometry to record daily FVC, symptoms, and other health outcomes is acceptable and feasible for patients with PF. Patients with PF show willingness to record home-based spirometry data regularly without prompting, which suggests that they are interested in monitoring their lung function. Some are willing to continue recording home-based spirometry and other outcomes on a long-term basis even when not involved in a formal clinical trial or survey. When prompts to record other health outcomes (eg, PROM) are added, these prompts have resulted in regular reporting of those outcomes. It is anticipated that use of prompts or triggers to record

spirometry will result in more sustained and frequent collection of home-based spirometry data.

This study was led by patient advocacy groups, and all participants used the patientMpower app independently without help from health care professionals. Our observations suggest that using a mobile device-based app linked to appropriate sensor devices is feasible to capture multiple long-term patient-reported objective and subjective outcomes within a single platform for studies in pulmonary fibrosis. This work was undertaken well ahead of the COVID-19 pandemic announced in March 2020. Our findings have informed the longitudinal clinical studies on the home-based monitoring of patients with fibrotic lung diseases to guide medication adjustments and nonpharmacological support strategies.

Acknowledgments

We are very grateful to all of the participants from the Irish Lung Fibrosis Association and PF Warriors patient groups for testing the application and providing their data. We are particularly grateful to the participant who allowed us to present their data in Figure 2. Gemma O'Dowd of ILFA assisted in participant recruitment and collection of data in Ireland. Data listings were prepared by Martin Holmes and Rebecca Neish, and writing assistance was provided by Amy Boulstridge (all of patientMpower Ltd).

This study was funded by patientMpower Ltd.

Authors' Contributions

CE and EC conceived the design of the studies. NC, BV, CE, and AMR designed and approved the study protocol and participant information materials. EC developed the patientMpower electronic health journal. NC and BV communicated the study to their respective patient advocacy groups. CE and AMR conducted the data analysis and drafted this manuscript. All authors participated in different parts of the concept, design, and data collection and entry; revised the article critically for important intellectual content; and gave their final approval of the version to be published.

Conflicts of Interest

CE and EC are employees and shareholders of patientMpower Ltd. BV is an adviser to patientMpower Ltd (reimbursed).

Multimedia Appendix 1
Description of the participant questionnaire.
[DOC File , 31 KB - mhealth_v8i11e16158_app1.doc]

References

1. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, ATS/ERS/JRS/ALAT Committee on Idiopathic Pulmonary Fibrosis. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. Am J Respir Crit Care Med 2011 Mar 15;183(6):788-824 [FREE Full text] [doi: 10.1164/rccm.2009-040GL] [Medline: 21471066]
2. Collard HR, Moore BB, Flaherty KR, Brown KK, Kaner RJ, King TE, Idiopathic Pulmonary Fibrosis Clinical Research Network Investigators. Acute exacerbations of idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2007 Oct 01;176(7):636-643 [FREE Full text] [doi: 10.1164/rccm.200703-463PP] [Medline: 17585107]
3. Richeldi L, Collard HR, Jones MG. Idiopathic pulmonary fibrosis. Lancet 2017 May 13;389(10082):1941-1952. [doi: 10.1016/S0140-6736(17)30866-8] [Medline: 28365056]
4. National Institute for Health and Care Excellence. Idiopathic Pulmonary Fibrosis in Adults: Diagnosis and Management. Clinical Guideline CG 163. London: National Institute for Health and Care Excellence; 2017 May 23. URL: https://www.nice.org.uk/guidance/cg163 [accessed 2019-09-01]
5. Russell A, Jones G, Sabella R, Cassidy N, Burge G, Lines S. Health status and impact of living with idiopathic pulmonary fibrosis (IPF): UK and Ireland Delphi survey. In: Eur Resp J 2016, 48 (suppl 60) OA4573. 2016 Sep Presented at: European Respiratory Society; September 2016; London. [doi: 10.1183/13993003.congress-2016.oa4573]
6. Russell A, Adamali H, Molyneaux PL, Lukey PT, Marshall RP, Renzoni EA, et al. Daily Home Spirometry: An Effective Tool for Detecting Progression in Idiopathic Pulmonary Fibrosis. Am J Respir Crit Care Med 2016 Oct 15;194(8):989-997 [FREE Full text] [doi: 10.1164/rccm.201511-2152OC] [Medline: 27089018]
7. Johannson KA, Vittinghoff E, Morisset J, Lee JS, Balmes J, Collard HR. Home monitoring improves endpoint efficiency in idiopathic pulmonary fibrosis. Eur Respir J 2017 Jul;50(1) [FREE Full text] [doi: 10.1183/13993003.02406-2016] [Medline: 28679608]
8. Maher T, Cottin V, Russell A, Corte T, Hammer P, Michael A. Correlation between home and clinic spirometry in subjects with IPF: results from the INMARK trial. In: Eur Respir J. 2019 Sep Presented at: European Respiratory Society; 2019; Madrid URL: https://erj.ersjournals.com/content/54/suppl_63/PA1318 [doi: 10.1183/13993003.congress-2019.PA1318]

9. Edwards C, Cassidy N, Costello E, Russell A. Pilot Study to Test the Utility and Acceptability of An Electronic Health Record (patientMpower) for Patients With Lung Fibrosis. In: iproc. 2017 Sep 22 Presented at: Connected Health; 2017; Boston p. e10 URL: https://www.iproc.org/2017/1/e10/ [doi: 10.2196/iproc.8389]
10. Edwards C, Vick B, Costello E. User experience of patientMpower electronic health journal for pulmonary fibrosis in volunteers recruited via the PF Warriors support group. In: Am J Respir Crit Care Med. 2018 Presented at: American Thoracic Society; 2018; San Diego p. A4932 URL: https://www.atsjournals.org/doi/10.1164/ajrccm-conference.2019.199.1_MeetingAbstracts.A4627 [doi: 10.1183/13993003.congress-2018.pa4810]
11. Goodyear-Smith F, Jackson C, Greenhalgh T. Co-design and implementation research: challenges and solutions for ethics committees. BMC Med Ethics 2015 Nov 16;16:78 [FREE Full text] [doi: 10.1186/s12910-015-0072-2] [Medline: 26573410]
12. Costello E. patientMpower for pulmonary fibrosis YouTube channel. YouTube patientMpower channel. 2017. URL: https://www.youtube.com/channel/UCIIxdd59yipDjt9XdOvw4Dw [accessed 2019-09-01]
13. Moor CC, van Manen MJ, Tak NC, van Noort E, Wijsenbeek MS. Development and feasibility of an eHealth tool for idiopathic pulmonary fibrosis. Eur Respir J 2018 Mar;51(3) [FREE Full text] [doi: 10.1183/13993003.02508-2017] [Medline: 29348156]
14. Moor CC, Wapenaar M, Miedema JR, Geelhoed JJM, Chandoesing PP, Wijsenbeek MS. A home monitoring program including real-time wireless home spirometry in idiopathic pulmonary fibrosis: a pilot study on experiences and barriers. Respir Res 2018 May 29;19(1):105 [FREE Full text] [doi: 10.1186/s12931-018-0810-3] [Medline: 29843728]
15. Walsh S, Cahill T, Edwards C, Costello E, Walsh J, Russell A, et al. Patient-reported monitoring of symptoms and spirometry via the patientMpower platform in idiopathic pulmonary fibrosis. In: Am J Respir Crit Care Med. 2018 Presented at: American Thoracic Society; 2018; San Diego p. A4933 URL: https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2018.197.1_MeetingAbstracts.A4933
16. Khan F, Stewart I, Howard L, McKeever TM, Jones S, Hearson G, et al. The Its Not JUST Idiopathic pulmonary fibrosis Study (INJUSTIS): description of the protocol for a multicentre prospective observational cohort study identifying biomarkers of progressive fibrotic lung disease. BMJ Open Respir Res 2019;6(1):e000439 [FREE Full text] [doi: 10.1136/bmjresp-2019-000439] [Medline: 31258922]
17. Fox L, McLeod L, Russell A, Edwards C, O'Reilly K. An observational study of spirometry and symptoms via a mobile electronic health journal in usual care of patients with idiopathic pulmonary fibrosis. In: Am J Respir Crit Care Med. 2020 Presented at: American Thoracic Society; 2020; Virtual p. A3394. [doi: 10.1164/ajrccm-conference.2020.201.1_meetingabstracts.a3394]
18. O'Carroll O, MacCann R, O'Reilly A, Dunican EM, Feeney ER, Ryan S, et al. Remote monitoring of oxygen saturation in individuals with COVID-19 pneumonia. Eur Respir J 2020 Aug;56(2) [FREE Full text] [doi: 10.1183/13993003.01492-2020] [Medline: 32616588]

Abbreviations

FVC: forced vital capacity
ILFA: Irish Lung Fibrosis Association
IPF: idiopathic pulmonary fibrosis
IPF-PROM: IPF patient-reported outcome measure
mMRC: modified Medical Research Council
PF: pulmonary fibrosis
PROM: patient-reported outcome measure

Edited by G Eysenbach, R Kukafka; submitted 06.09.19; peer-reviewed by A Younas, M Dahl; comments to author 23.09.20; revised version received 06.10.20; accepted 28.10.20; published 20.11.20.

Please cite as:
Edwards C, Costello E, Cassidy N, Vick B, Russell AM
Use of the patientMpower App With Home-Based Spirometry to Monitor the Symptoms and Impact of Fibrotic Lung Conditions: Longitudinal Observational Study
JMIR Mhealth Uhealth 2020;8(11):e16158
URL: http://mhealth.jmir.org/2020/11/e16158/
doi:10.2196/16158
PMID:33216007

Original Paper

Using Mobile Health Technology to Deliver a Community-Based Closed-Loop Management System for Chronic Obstructive Pulmonary Disease Patients in Remote Areas of China: Development and Prospective Observational Study

Ning Deng[1,2*], PhD; Juan Chen[3*], MD; Yiyuan Liu[1], MSc; Shuoshuo Wei[3], MSc; Leiyi Sheng[1], BSc; Rong Lu[3], BSc; Zheyu Wang[1], BSc; Jiarong Zhu[3], BSc; Jiye An[1], PhD; Bei Wang[3], BSc; Hui Lin[1], BSc; Xiuyan Wang[3], MSc; Yumin Zhou[4], MD; Huilong Duan[1], PhD; Pixin Ran[4], MD

[1]Ministry of Education Key Laboratory of Biomedical Engineering, College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China

[2]Engineering Research Center of Cognitive Healthcare of Zhejiang Province (Sir Run Run Shaw Hospital), Zhejiang University, Hangzhou, China

[3]Department of Pulmonary and Critical Care Medicine, General Hospital of Ningxia Medical University, Yinchuan, China

[4]State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Diseases, Guangzhou Institute of Respiratory Health, First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Pixin Ran, MD
State Key Laboratory of Respiratory Disease
National Clinical Research Center for Respiratory Diseases, Guangzhou Institute of Respiratory Health
First Affiliated Hospital of Guangzhou Medical University
195 Dongfeng Xi Road
Guangzhou,
China
Phone: 86 020 3710 3001
Fax: 86 020 3710 3001
Email: pxran@gzhmu.edu.cn

Abstract

Background: Mobile health (mHealth) technology is an increasingly recognized and effective method for disease management and has the potential to intervene in pulmonary function, exacerbation risk, and psychological status of patients with chronic obstructive pulmonary disease (COPD).

Objective: This study aimed to investigate the feasibility of an mHealth-based COPD management system designed for Chinese remote areas with many potential COPD patients but limited medical resources.

Methods: The system was implemented based on a tailored closed-loop care pathway that breaks the heavy management tasks into detailed pieces to be quantified and executed by computers. Low-cost COPD evaluation and questionnaire-based psychological intervention are the 2 main characteristics of the pathway. A 6-month prospective observational study at the community level was performed to evaluate the effect of the system. Primary outcomes included changes in peak expiratory flow values, quality of life measured using the COPD assessment test scale, and psychological condition. Acute exacerbations, compliance, and adverse events were also measured during the study. Compliance was defined as the ratio of the actual frequency of self-monitoring records to the prescribed number.

Results: A total of 56 patients was enrolled; 39 patients completed the 6-month study. There was no significant difference in the mean peak expiratory flow value before and after the 6-month period (366.1, SD 106.7 versus 313.1, SD 116.6; *P*=.11). Psychological condition significantly improved after 6 months, especially for depression, as measured using the Patient Health Questionnaire-9 scale (median 6.0, IQR 3.0-9.0 versus median 4.0, IQR 0.0-6.0; *P*=.001). The COPD assessment test score after 6 months of intervention was also lower than that at the baseline, and the difference was significant (median 4.0, IQR 1.0-6.0 versus median 3.0, IQR 0.0-6.0; *P*=.003). The median overall compliance was 91.1% (IQR 67%-100%). In terms of acute exacerbation, 110 exacerbations were detected and confirmed by health care providers (per 6 months, median 2.0, IQR 1.0-5.0).

Moreover, 72 adverse events occurred during the study, including 1 death, 19 hospitalizations, and 52 clinic visits due to persistent respiratory symptoms.

Conclusions: We designed and validated a feasible mHealth-based method to manage COPD in remote Chinese areas with limited medical resources. The proposed closed-loop care pathway was effective at the community level. Proper education and frequent communication with health care providers may encourage patients' acceptance and use of smartphones to support COPD self-management. In addition, WeChat might play an important role in improving patient compliance and psychological distress. Further research might explore the effect of such systems on a larger scale and at a higher evidence level.

(JMIR Mhealth Uhealth 2020;8(11):e15978) doi:10.2196/15978

KEYWORDS

COPD; mobile health technology; closed-loop care pathway; chronic disease management; exacerbations

Introduction

Chronic obstructive pulmonary disease (COPD) is one of the most prevalent chronic diseases in China. According to the latest national survey, approximately 99.9 million Chinese adults aged ≥20 years have spirometry-defined COPD [1]. Pulmonary function tests (PFTs) are essential for the diagnosis of COPD [2,3]. However, only 12% of Chinese patients with COPD can receive full PFTs [1], and even fewer can receive full PFTs in remote areas. For example, the Ningxia Hui Autonomous Region is a relatively remote area in China with a population of over 6 million people. The prevalence of COPD in Ningxia is 8.9% in people aged ≥40 years [4]. Only 23.6% of COPD cases are diagnosed, and only 23.3% of COPD cases are treated [4].

Moreover, most COPD patients are unaware of their condition, and few receive follow-ups from doctors. Once diagnosed with COPD, effective management should be performed based on an individualized assessment to improve prognosis [5], consisting of long-term self-management by patients and timely intervention from doctors [6]. Mobile health (mHealth) has become a promising tool for the management of chronic diseases by health care providers [7]. Furthermore, patients are willing to use smartphones to manage chronic diseases [8-10].

Several studies have explored the feasibility of an mHealth-based intervention for COPD management in patients during the stable phase. Farmer et al [11] performed a randomized controlled trial with an intervention group using an internet-linked, tablet computer–based system to manage COPD. van der Heijden et al [12] developed a mobile system to predict and detect exacerbation using an oximeter and spirometer. Zhang et al [13] implemented a smartphone-based Internet of Things system that provided medication reminders, data collection, health education, and communication for COPD self-management. This project offered PFT in community centers in Shanghai. These studies indicate that mHealth systems have the potential to help patients effectively manage COPD. However, most of these studies assumed that patients were in an ideal environment with portable devices and abundant community health care resources, which may not be applicable in remote Chinese areas.

In areas with decreased resources such as Ningxia that have many potential COPD patients but not enough medical resources, the aforementioned pattern of mHealth-based interventions may not work. First, spirometry is a complex and not readily available test in Chinese remote areas. COPD patients in these areas need a simpler and low-cost approach to help monitor pulmonary function. Second, health care providers in remote areas might not perform a guideline-based, comprehensive intervention, which results in a significant gap between standards of care and medical practice [14]. Third, during long-term management, depression and anxiety are common in COPD patients [15-17] and are strongly related to rehospitalization and exacerbation [18-20]. Therefore, appropriate psychological intervention is an essential preventive strategy. Moreover, COPD patients in remote areas have poor knowledge about their disease, thus leading to low self-management ability. Health care providers need to communicate more with patients to assess their condition and guide treatment.

To address these problems, we propose a pilot closed-loop care pathway that transforms the core content of long-term COPD management into a form that can be recognized and executed by the computer. The 2 main characteristics of the care pathway are (1) COPD evaluation combining peak expiratory flow (PEF) with the COPD assessment test (CAT) and (2) psychological intervention based on multiple scales. Our primary objective was to determine whether the introduction of an mHealth system based on the proposed care pathway could improve the COPD-specific health status of patients, such as pulmonary function, quality of life (QoL), and psychological conditions, in remote areas of China. Our secondary objectives were to determine if there were changes in exacerbations, hospitalizations, or clinic visits after implementation of the mobile app.

Methods

Study Design and Patient Recruitment

We conducted a 6-month prospective observational study to explore the effect of the system in a community-based primary care setting. The study was performed in Yinchuan, Ningxia Hui Autonomous Region, China. Patient recruitment was performed between September 2017 and June 2018. The patients were recruited from outpatient clinics of the General Hospital of Ningxia Medical University and its affiliated hospitals throughout Ningxia. Inclusion criteria included (1) age ≥40 years, (2) confirmed diagnosis of COPD according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines [5], and (3) access to a smartphone. Exclusion criteria

included (1) diagnosis of other lung diseases with similar symptoms or any other disease-causing changes in pulmonary function, (2) significant mental health disorder, (3) visual or aural disorders, (4) no home internet access, and (5) involvement in another study.

Patients meeting the criteria and willing to participate in the study signed the informed consent forms. Ethical approval was granted by the Ethics Committee for the Conduct of Human Research at the General Hospital of Ningxia Medical University (NXMU-GH-2017-273). Baseline data, including demographics, clinical characteristics, and primary outcomes, were first collected through laboratory examination and face-to-face interviews. Patients were instructed on how to install and use the app at the same time. Moreover, we provided a peak flow meter for each patient and ensured that they could correctly use it themselves.

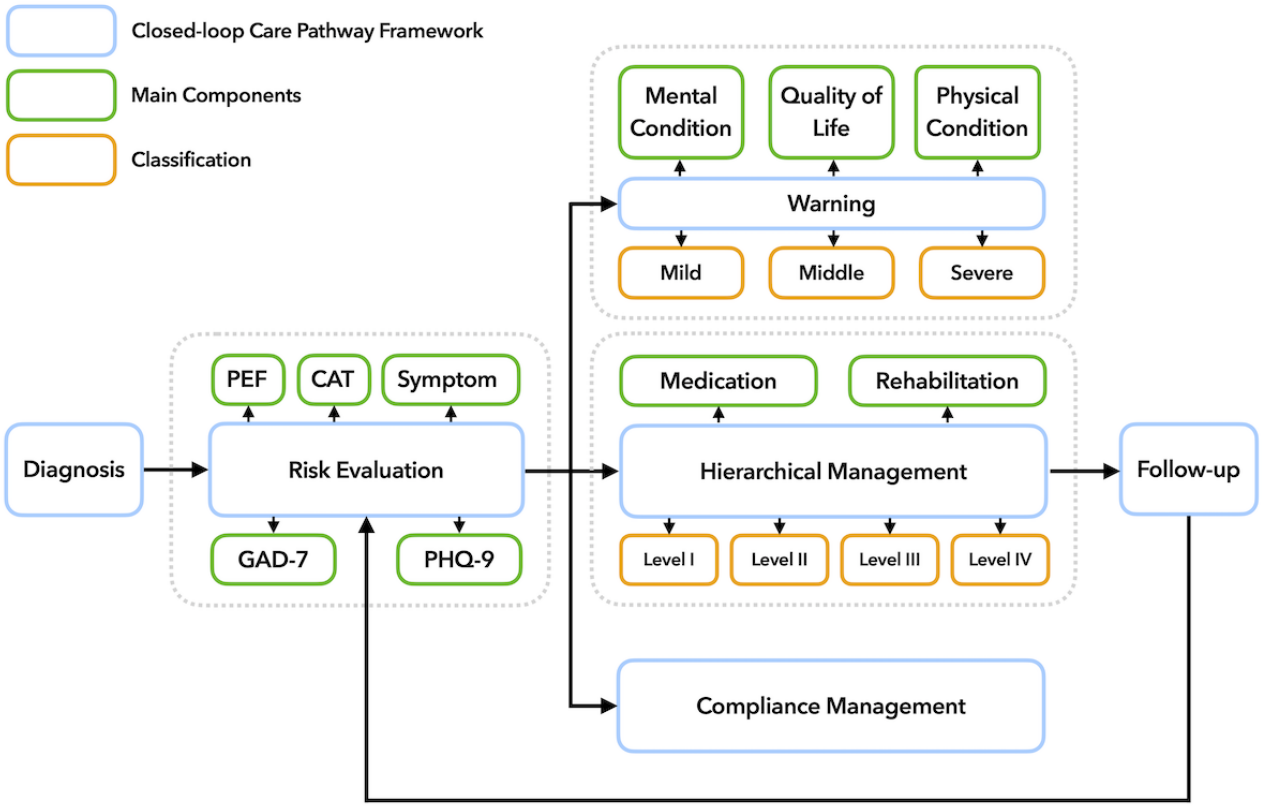
After enrollment, patients started self-management using the app, and health care providers supervised patients through the website. In this study, the health care providers consisted of physicians from the hospital and general practitioners from community health centers. General practitioners were responsible for daily management, whereas physicians were responsible for the treatment of acute exacerbations and professional support. All management procedures followed the proposed closed-loop care pathway. Follow-up was the main intervention for health care providers. Generally, follow-up was conducted every 2 weeks. In the event of a warning or low compliance, an additional follow-up was arranged. Telephone was the traditional tool for follow-up; however, sometimes it is inefficient because patients may not be able to conveniently answer the call or may not want to be disturbed. Therefore, we introduced WeChat, the most popular instant messaging app in China, as an alternative follow-up method for patients.

Furthermore, a WeChat chat group including all patients and 2 health care providers was formed during the study to facilitate communication. Patients needed to login to WeChat with a password, and they could only be invited by the doctors to join the group chat. Only common questions with no sensitive medical information were included in the group. Patients could chat privately with the doctor on WeChat regarding private information.

Closed-Loop Care Pathway Design

The care pathway mainly included risk evaluation, hierarchical management, follow-up, warning, and compliance management, as shown in Figure 1. Following this pathway, patients were first evaluated and classified into different levels, and then different management strategies including medication and pulmonary rehabilitation were provided. Health care providers performed regular follow-ups with patients to track their health status. Patients' daily self-monitoring data and compliance were evaluated as well to detect abnormal conditions. A detailed description of each module can be found in Multimedia Appendix 1.

Figure 1. Flow diagram of the proposed care pathway. CAT: Chronic Obstructive Pulmonary Disease Assessment Test; GAD-7: Generalized Anxiety Disorder-7; PEF: peak expiratory flow; PHQ-9: Patient Health Questionnaire-9.

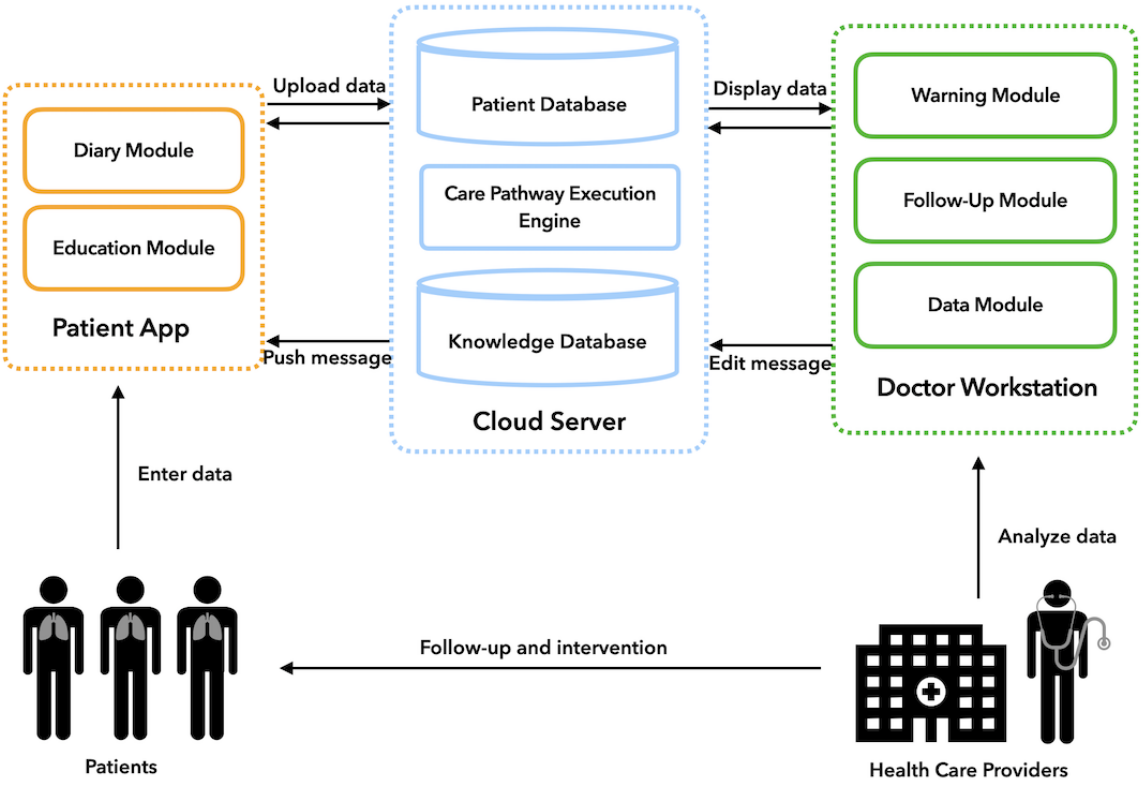


Because of the complexity of spirometry and the lack of professional portable device support in Chinese remote areas, we chose a simpler indicator suitable for home-based management, the PEF, which is the maximum flow achieved during an expiration delivered with maximum force starting from the level of maximal lung inflation [21]. Studies have shown that a PEF rate <80% of predicted may be a good indicator for detecting patients with COPD [22]. Increased variability in PEF could be considered an early index of COPD development [23]. Furthermore, COPD is known to impact patients beyond just dyspnea [5]. Therefore, the CAT, a simple and validated instrument to assess the impact of COPD in routine life [24], is strongly recommended by GOLD. In addition, to monitor patients' psychological state, patients completed the Patient Health Questionnaire-9 (PHQ-9) scale [25] for depression and Generalized Anxiety Disorder-7 (GAD-7) scale [26] for anxiety. The participants were asked to log their PEF value (the highest value of 3 repeated measurements), CAT, GAD-7, and PHQ-9 scales once every 2 weeks over the 6-month period.

mHealth System Description

Based on this pathway, we implemented our mHealth-based COPD management system. As shown in Figure 2, the system contained 3 parts: patient app, doctor workstation, and cloud server. The patient app ran on a smartphone and had 2 main modules: diary module and education module. The diary module guided patients to record daily medication and symptoms. Moreover, every 2 weeks, patients were required to record their PEF value, CAT, PHQ-9, and GAD-7 scales through this module. The PEF value can be measured via a simple peak flow meter, which is available in remote Chinese areas. The education module included COPD-related educational materials in the form of videos and text. Patients received personalized materials according to their current situation [27]. We provided 2 versions of the app (Android and iOS). Both versions had the same functionality, and all the data were synchronized with the cloud.

Figure 2. System architecture.



The doctor workstation was a web-based platform for health care providers with 3 modules: warning module, follow-up module, and data module. The warning module alerted health care providers as soon as the system detected abnormal data uploaded by patients. The follow-up module provided a template for health care providers to record each follow-up in a uniform format. The data module provided complete patient information, including demographics, in-hospital examination results, and self-monitoring data history. Detailed screenshots and a function description of the patient app and doctor workstation can be found in Multimedia Appendix 2.

The cloud server securely stored all data and had an inference engine to provide decision support to the clients based on the pathway, such as patient classification, follow-up scheduling, and warning generation. In addition, it had a recommendation service for delivering educational material.

Data Collection and Analysis

Primary and Secondary Outcomes

The primary outcomes included changes in PEF values; changes in QoL, as measured by the CAT scale; and changes in psychological condition, as measured by the PHQ-9 and GAD-7 scales. These outcomes reflected the COPD-specific health

status of patients. As mentioned before, PEF values were measured using a peak flow meter and were uploaded via the patient app, while the 3 scales (CAT, PHQ-9, and GAD-7) were completed directly in the patient app.

The secondary outcome was the number of acute exacerbations during the study. An acute exacerbation was classified as a type of warning in the pathway and was detected from symptoms recorded by the patients. We considered dyspnea, sputum purulence, and sputum volume as major symptoms, each scored as 5 points, and nasal discharge or congestion, sore throat, cough, and wheeze as minor symptoms, each scored as 1 point, according to a previously validated definition [28]. A suspected exacerbation was automatically detected by the system if the summed symptom score was >6 points for ≥2 consecutive days. The detection required further confirmation by health care providers to ensure that the symptoms were caused by COPD.

Compliance

In addition to the outcomes, patient compliance during the study was analyzed because of its relevance to self-efficacy [29] and its potential to show system availability. In this study, we defined compliance as the ratio of actual frequency of records to the prescribed number of records for the PEF, CAT scale, PHQ-9 scale, and GAD-7 scale [26], as shown in Figure 3. Patients were required to upload these 4 types of data every 2 weeks. We calculated the compliance for each month within the 6-month period.

Figure 3. Calculation of patient compliance during a fixed period.

$$compliance = \min\left(\frac{N(\mathrm{PEF}) + \mathrm{N(CAT)} + \mathrm{N(PHQ)} + \mathrm{N(GAD)}}{N(\text{prescribed number of records})^{\#}} \times 100\%, 1\right)$$

N = 4 for a period of 2 weeks

Adverse Events

Information regarding serious adverse events was collected and recorded for further analysis. These events included deaths, hospitalizations, and clinic visits due to persistent respiratory symptoms. Only records related to COPD were counted. The adverse outcomes were tracked mainly in 2 ways. First, health care providers checked the hospital information system of the General Hospital of Ningxia Medical University and its affiliated hospitals regularly for records of hospital admission and clinic visits. Second, when conducting routine follow-ups, patients were asked whether they had gone to other hospitals during the study period. For the clinic visits and hospitalizations recorded for adverse outcomes, we also collected records of patients during the 6 months prior to enrollment in the study.

Statistical Analysis

Python 3.6 was used for data preprocessing, including extraction of clinical outcomes from the database and calculation of compliance. SPSS V23.0 was used for statistical analyses. A paired Student *t* test was used to analyze changes in PEF values. The original scores of the 3 scales before and after the 6-month study were analyzed using Wilcoxon signed-rank tests. A *P* value <.05 was considered statistically significant. Table 1 summarizes the data collected in this study.

Table 1. Data collection and statistical analysis.

Data	Collection time	Collection method	Statistical test
Primary clinical outcomes			
PEF[a]	At baseline and after 6 months	Peak flow meter	Paired *t* test
CAT[b]	At baseline and after 6 months	Scale	Wilcoxon signed-rank test
PHQ-9[c]	At baseline and after 6 months	Scale	Wilcoxon signed-rank test
GAD-7[d]	At baseline and after 6 months	Scale	Wilcoxon signed-rank test
Secondary clinical outcome			
Acute exacerbation	During the study period	Detected by the system according to symptoms and confirmed by health care providers	None
Compliance	During the study period (each month)	Calculated by the system	None
Adverse event			
Death	During the study period	From HIS[e] and follow-up	None
Hospitalizations	During the 6 months prior to enrollment and during the study period	From HIS and follow-up	None
Clinic visits	During the 6 months prior to enrollment and during the study period	From HIS and follow-up	None

[a]PEF: peak expiratory flow.

[b]CAT: Chronic obstructive pulmonary disease Assessment Test.

[c]GAD-7: Generalized Anxiety Disorder 7.

[d]PHQ-9: Patient Health Questionnaire 9.

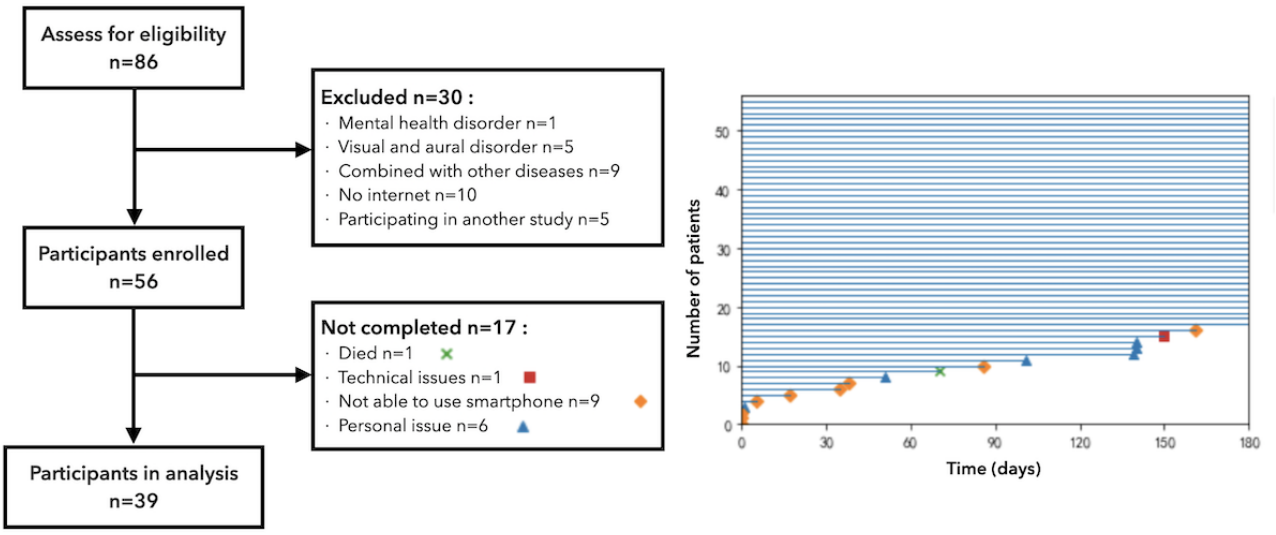
[e]HIS: hospital information system.

Results

Participant Characteristics

A total of 56 patients was enrolled in the study, and 39 patients completed the study. Seventeen participants (17/56, 30%) quit the study prematurely for various reasons. Figure 4 shows the flow diagram of patient recruitment and the study timeline that demonstrates the exit point of the 17 participants.

Figure 4. Flow diagram of patient recruitment and the study timeline.



The demographics and clinical characteristics of the 39 participants who completed the study are presented in Table 2. There were more male patients than female patients (male: 36/39, 92%; female: 3/39, 8%), and the age ranged from 49

years old to 76 years old. Approximately 85% (33/39) of the participants had an educational background of high school or below, and only 15% (6/39) had at least a college degree. A total of 33 participants were ex-smokers (30/39, 77%) or current smokers (3/39, 8%). Furthermore, the 3 current smokers quit smoking after the study. In terms of COPD severity and pulmonary function, 18% (7/39), 41% (16/39), 15% (6/39), and 26% (10/39) of the participants were at the GOLD1, GOLD2, GOLD3, or GOLD4 stage, respectively. The mean baseline forced expiratory volume in 1 second was 49.91% of predicted (95% CI, 42.54%-57.27%).

Table 2. Characteristics of patients who completed the study (n=39).

Patient demographic and clinical characteristics	Results
Sex, n (%)	
Male	36 (92)
Female	3 (8)
Age (years), mean (SD)	61.82 (6.09)
Education background, n (%)	
Primary school and below	5 (13)
Secondary school	12 (31)
High school	16 (41)
Graduate and above	6 (15)
Smoking, n (%)	
Ex-smoker	30 (77)
Current smoker	3 (8)
Non-smoker	6 (15)
Pulmonary function	
FEV1[a] %pred (%), mean (SD)	49.91 (22.73)
FEV1/FVC[b] (%), mean (SD)	51.45 (13.97)
COPD[c] severity, n (%)	
GOLD1[d]	7 (18)
GOLD2	16 (41)
GOLD3	6 (15)
GOLD4	10 (26)

[a]FEV1: forced expiratory volume in 1 second.

[b]FVC: forced vital capacity.

[c]COPD: chronic obstructive pulmonary disease.

[d]GOLD: The Global Initiative for Obstructive Lung Disease.

Primary and Secondary Outcomes

Table 3 presents the clinical outcomes of the study: PEF for airflow obstruction, CAT for QoL, and PHQ-9 and GAD-7 for psychological state. The psychological state of patients who completed the study significantly improved after the 6-month intervention, especially for depression, as measured by the PHQ-9 scale (P=.001). The CAT score after 6 months of management was also significantly lower than that at baseline (P=.003). In terms of PEF, the mean value decreased after management but was not significant (P=.11). From a clinical viewpoint, the median CAT score remained at a medium level (11-20 [5]), and the median GAD-7 score remained at the “no anxiety disorder” level (0-4 [26]) before and after the study. The median PHQ-9 score changed from mild depression (5-9 [25,30]) to minimal depression (0-4 [25]).

Table 3. Clinical outcomes before and after the study for patients who completed the study (n=39).

Outcomes	Baseline	After 6 months	*P* value
PEF[a], mean (SD)	366.1 (106.7)	313.1 (116.6)	.11
CAT[b], median (Q25[c]-Q75[d])	17.0 (14.0-23.0)	14 (10.0-18.0)	.003
PHQ-9[e], median (Q25-Q75)	6.0 (3.0-9.0)	4.0 (0.0-6.0)	.001
GAD-7[f], median (Q25-Q75)	4.0 (1.0-6.0)	3.0 (0.0-6.0)	.01

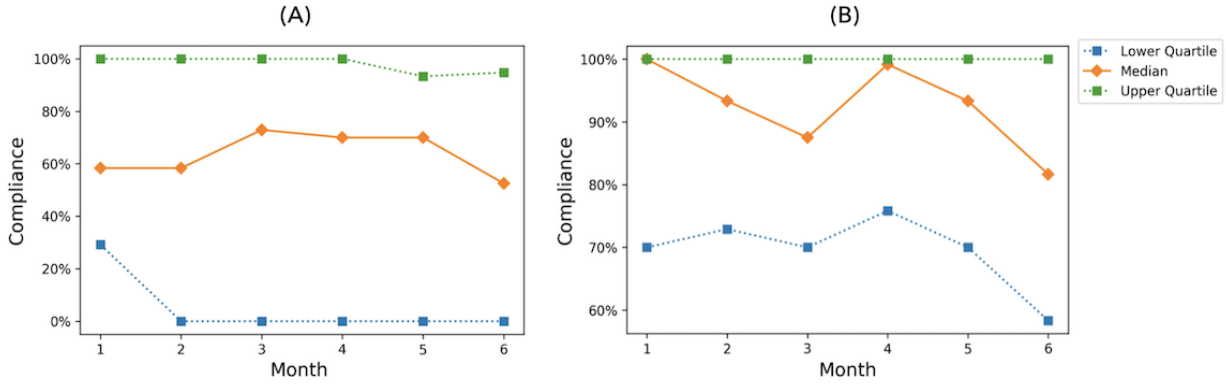
[a]PEF: peak expiratory flow.
[b]CAT: COPD Assessment Test.
[c]Q25: first quartile.
[d]Q75: third quartile.
[e]PHQ-9: Patient Health Questionnaire 9.
[f]GAD-7: Generalized Anxiety Disorder 7.

In terms of acute exacerbations, 459 warning events were detected by our system during the study, 117 of which were acute exacerbation warnings related to symptoms. Among these symptom warning events, 7 were excluded due to worsening symptoms caused by cardiovascular disease and other diseases, and only 110 events were confirmed as exacerbation of COPD by health care providers (median 2.0 per 6 months, IQR 1.0-5.0). Moreover, 64.5% (71/110) of these exacerbations ended up as a clinic visit or hospitalization. The detailed number of acute exacerbations for each patient can be found in Multimedia Appendix 3.

Compliance

Figure 5 shows the participant compliance per month for all enrolled patients (n=56) and patients who completed the study (n=39). Some participants recorded data more often than prescribed; thus, their compliance was over 100% but was calculated as 100%. Because more than one-quarter of participants withdrew midway, the lower quartile in Figure 5A tended to be 0%. For those who completed the study (n=39), monthly compliance remained at a high level, as shown in Figure 5B. We also calculated the overall compliance of these patients during the 6-month period. The median was 91.1%, (IQR 67%-100%). The overall trends of the uploaded data can also be found in Multimedia Appendix 3.

Figure 5. Quartiles of and median participant compliance per month over the 6-month period for (A) all enrolled patients (n=56) and (B) those who completed the study (n=39).



Adverse Events

According to our definition, a total of 72 adverse events occurred during the study, including 1 death, 52 clinic visits, and 19 hospitalizations. We compared the number of hospitalizations and clinic visits for each patient before and after the study (see Multimedia Appendix 3 for detailed results). During the 6 months prior to enrollment, patients (n=39) reported a total of 115 clinic visits and 69 hospitalizations related to COPD. Both clinic visits (decreased of 54.8%) and hospitalizations (decrease of 72.5%) decreased over the 6-month period compared with the same period before the study.

Discussion

Principal Findings

In this study, we designed and implemented a feasible method to manage COPD in Chinese remote areas using an mHealth-based system over a 6-month period, despite the challenges with pulmonary function monitoring outside of the hospital. Although pulmonary function did not significantly change with the implementation of this system, we found an improvement in COPD-specific QoL and mental health assessments of participants.

Some of our results could be attributed to the diversity of COPD severity among participants. In terms of acute exacerbation, a total of 110 exacerbations w identified during the study, because COPD severity for >80% of participants was at least a GOLD2 level. The median number of exacerbations (2.0 per 6 months) was close to the result (3.0 per year) from a previous survey conducted in the Asia-Pacific region [31]. According to the records of adverse events, over half of these exacerbation cases led to a clinic visit or hospitalization. The rest of the exacerbations were handled remotely by health care providers. The high variability in COPD severity also explains the large SDs of the PEF values.

Of the 56 patients, 17 patients quit midway through the study, mainly due to difficulty using smartphones. The relatively high dropout rate could be attributed to the low educational attainment of participants (about 85% had achieved an educational level of high school or below). For patients who completed the study (n=39), the overall median 6-month compliance reached 91.1%, which was relatively high compared with the rates in prior studies [11,32]. In addition, after the 6-month intervention, 3 current smokers quit smoking. These results indicate that, despite a low educational level, our pathway-driven, timely intervention played an important role in behavioral change by COPD patients.

Comparison With Prior Work

To understand the innovation of this study, we compared our study with prior work from multiple aspects. In terms of QoL, significant improvement in COPD-specific QoL was observed in this study, whereas prior studies reported no change in QoL or only changes in general health status [11,32-35]. CAT and EuroQol-5D (EQ-5D) [36] are the most popular scales for COPD management [11,32-35]. CAT is more disease-specific, whereas EQ-5D focuses more on general health status. Studies using the EQ-5D for QoL measurement often introduced other scales as a supplement for COPD-specific symptoms [11,32]. More scales capture more details for research but place an extra burden on patients, which may lead to poor compliance. Moreover, previous research has shown that the CAT is more suitable for the comprehensive assessment of COPD severity [37]. Therefore, we chose the CAT as the only outcome for QoL.

In terms of pulmonary function, prior studies [33,37,38] did not directly compare changes in pulmonary function measured via portable devices at home. In this study, we chose PEF for continuous pulmonary function monitoring, and the PEF values showed a downward trend because of the irreversible decline of pulmonary function. Although portable spirometers are available and have been used in some studies [33,35,38], mass promotion is difficult in China considering the economic and population factors. Equipped with a peak flow meter and an mHealth system, PEF is easy to measure at home and under supervision. Moreover, to some degree, PEF measurement alone cannot be reliably used as the only diagnostic test because of its weak specificity [5]; hence, we combined PEF with CAT and symptoms to evaluate COPD-specific health status and monitor acute exacerbations. Our study showed that elderly people with a low educational background could master PEF measurements. With limited medical resources, PEF measured by a peak flow meter is suitable for COPD management in Chinese remote areas.

In terms of psychological conditions, except for our study, only 1 study [34] found significant improvement in psychological distress with a long-term (12-month) telehealth intervention but without a clear explanation. Similar to the studies by Farmer et al [11] and Rixon et al [34], we used independent psychological scales, whereas other studies investigated psychological factors within general health status such as with the EQ-5D. In this study, the improvement in psychological status may be explained by the high frequency of communication between patients and health care providers. The interval of routine follow-up visits in the study by Farmer et al [11] was 3 months, and extra contact only occurred for safety alerts, whereas in our study, follow-ups were arranged at least every 2 weeks. With the assistance of WeChat, the communication frequency was even higher in practice.

In terms of the intervention, the care pathway was a unique strategy that we proposed in the field of COPD management, breaking the heavy management tasks into detailed pieces that can be quantified and executed by computers. Combined with mHealth, it was simple and convenient to deliver the COPD management system driven by the care pathway. With the help of this system, both patients and health care providers were aware of what exactly to do in different phases of management. Although the pathway is not yet perfect, it could improve the outcome of patients and efficiency of health care providers.

Strengths and Limitations

Our study has several strengths. First, we defined compliance management as part of the care pathway. An extra follow-up would be required for patients with local compliance. In this study, inspired by our previous study [39], we calculated patient compliance according to the frequency of their self-monitoring behaviors. Compared with prior studies [11,32], our compliance was more objective and practical over the long term. According to our results, the monthly compliance remained high for patients who completed the study. Second, WeChat played a surprisingly important role in our study. WeChat was first designed as a follow-up tool, but as the study progressed, patients found it a convenient way of communication and to have their questions answered. The WeChat chat group produced dozens of messages every day during the study. Patients said that they could learn a lot about self-management skills in the group by just viewing messages because many common questions were raised and answered immediately. To some extent, getting feedback in a short time helped patients build trust with health care providers and maintain high compliance. From another point of view, group talking introduced social activity for COPD patients, and doctors could intervene psychologically by managing patients' mental health and encouraging group members. This is a convincing reason for the mental health improvement in our study. The impact of social apps on COPD management may be worthy of further study.

Numerous potential weaknesses need to be acknowledged. First, the study was limited by the sample size and gender imbalance. The number of participants was relatively small (n=39), while

92% of the participants were male. The gender imbalance was mainly due to the higher incidence of COPD in men than in women in Ningxia [4]. The effect of our system on a larger scale and female patients needs further investigation. Second, when collecting the records from clinic visits and hospitalizations, an inherent bias existed because patients might go to hospitals for all related symptoms prior to the study due to a lack of communication with health care providers. Therefore, the difference in pre-study and post-study records may not reflect the change in disease condition. Moreover, from the methodological perspective, our pilot observational study could not yield the highest level of evidence, and a randomized controlled trial should be performed for further study.

Future Work

In future work, we will include more female patients and conduct a high evidence-level clinical trial. A comparison test should be conducted to determine if our intervention can perform well on a larger scale. We also plan to investigate the actual effect of WeChat in improving patient outcomes through the trial. The satisfaction rate of patients using WeChat may be presented as an outcome. Another direction for future work is to explore a more intelligent care pathway to provide more individualized and precise care for patients with COPD. The feasibility of applying a data-driven artificial intelligence approach to strengthen the current pathway will be investigated.

Conclusions

We introduced an mHealth-based method for COPD management at the community level. The tailored closed-loop care pathway was found to be feasible and effective in areas with limited medical resources. Although the decrease in pulmonary function was irreversible, patients' COPD-specific QoL and psychological status improved significantly after our 6-month study. Moreover, despite the relatively low educational background of participants, proper education and frequent communication may encourage their acceptance and use of smartphones to support COPD self-management. Researchers should choose the most appropriate intervention strategy according to the actual situation of the intervention area.

Acknowledgments

This study was supported by the National Key Research and Development Programs of China (No. 2017YFC0114105, 2016YFC0901703, 2018YFC0910503); National Scientific Foundation of China (No. 81760004); Open Project of State Key Laboratory of Respiratory Disease, China (No. SKLRD-OP-201903); Key research and development programs of Ningxia hui autonomous (No. 2018BFG02009, 2018BEG03035); and Open Fund of Engineering Research Center of Cognitive Healthcare of Zhejiang Province (Sir Run Run Shaw Hospital; No. 2018KFJJ02).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed description of the closed loop care pathway.
[PDF File (Adobe PDF File), 700 KB - mhealth_v8i11e15978_app1.pdf]

Multimedia Appendix 2
Detailed screenshots and function list of the mHealth system.
[PDF File (Adobe PDF File), 408 KB - mhealth_v8i11e15978_app2.pdf]

Multimedia Appendix 3
Detailed results of primary and secondary outcomes.
[PDF File (Adobe PDF File), 268 KB - mhealth_v8i11e15978_app3.pdf]

References

1. Wang C, Xu J, Yang L, Xu Y, Zhang X, Bai C, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study. The Lancet 2018 Apr 28;391(10131):1706-1717. [doi: 10.1016/S0140-6736(18)30841-9] [Medline: 29650248]
2. Tan WC, Bourbeau J, Aaron SD, Zhou G, Maltais F, Hernandez P, et al. Global Initiative for Chronic Obstructive Lung Disease 2017 Classification and Lung Function Decline in Chronic Obstructive Pulmonary Disease. Am J Respir Crit Care Med 2018 Mar 01;197(5):670-673. [doi: 10.1164/rccm.201706-1154LE] [Medline: 28858570]
3. British Medical Association. Investing in general practice: the new general medical services contract. 2003. URL: https://www.nhsemployers.org/~/media/Employers/Documents/SiteCollectionDocuments/gms_contract_cd_130209.pdf [accessed 2020-11-16]
4. Qiu J, Zhang Y, Chen J, Luo T, Yu X, Wang J, et al. [Prevalence of chronic obstructive pulmonary disease in Ningxia Hui Autonomous Region of China]. Zhonghua Jie He He Hu Xi Za Zhi 2013 Apr;36(4):265-268. [Medline: 23945339]

5. The Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis, Management and Prevention of Chronic Obstructive Pulmonary Disease: 2018 Report. 2018. URL: https://goldcopd.org/wp-content/uploads/2017/11/GOLD-2018-v6.0-FINAL-revised-20-Nov_WMS.pdf [accessed 2020-11-16]
6. Reynolds R, Dennis S, Hasan I, Slewa J, Chen W, Tian D, et al. A systematic review of chronic disease management interventions in primary care. BMC Fam Pract 2018 Jan 09;19(1):11 [FREE Full text] [doi: 10.1186/s12875-017-0692-3] [Medline: 29316889]
7. Wootton R. Twenty years of telemedicine in chronic disease management--an evidence synthesis. J Telemed Telecare 2012 Jun;18(4):211-220 [FREE Full text] [doi: 10.1258/jtt.2012.120219] [Medline: 22674020]
8. Dou K, Yu P, Deng N, Liu F, Guan Y, Li Z, et al. Patients' Acceptance of Smartphone Health Technology for Chronic Disease Management: A Theoretical Model and Empirical Test. JMIR Mhealth Uhealth 2017 Dec 06;5(12):e177 [FREE Full text] [doi: 10.2196/mhealth.7886] [Medline: 29212629]
9. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
10. Bengtsson U, Kjellgren K, Hallberg I, Lindwall M, Taft C. Improved Blood Pressure Control Using an Interactive Mobile Phone Support System. J Clin Hypertens (Greenwich) 2016 Feb;18(2):101-108 [FREE Full text] [doi: 10.1111/jch.12682] [Medline: 26456490]
11. Farmer A, Williams V, Velardo C, Shah SA, Yu L, Rutter H, et al. Self-Management Support Using a Digital Health System Compared With Usual Care for Chronic Obstructive Pulmonary Disease: Randomized Controlled Trial. J Med Internet Res 2017 May 03;19(5):e144 [FREE Full text] [doi: 10.2196/jmir.7116] [Medline: 28468749]
12. van der Heijden M, Lucas PJF, Lijnse B, Heijdra YF, Schermer TRJ. An autonomous mobile system for the management of COPD. J Biomed Inform 2013 Jun;46(3):458-469 [FREE Full text] [doi: 10.1016/j.jbi.2013.03.003] [Medline: 23500485]
13. Zhang J, Song Y, Bai C. MIOTIC study: a prospective, multicenter, randomized study to evaluate the long-term efficacy of mobile phone-based Internet of Things in the management of patients with stable COPD. Int J Chron Obstruct Pulmon Dis 2013;8:433-438 [FREE Full text] [doi: 10.2147/COPD.S50205] [Medline: 24082784]
14. Zhou XM, Wu C, Zhao L, Gao YZ, Yuan Y, Xiao XX, et al. [A cross-sectional survey of the knowledge on chronic obstructive pulmonary disease in physicians of tertiary hospitals in Northern China]. Zhonghua Nei Ke Za Zhi 2016 Sep 01;55(9):717-720. [doi: 10.3760/cma.j.issn.0578-1426.2016.09.012] [Medline: 27586981]
15. Maurer J, Rebbapragada V, Borson S, Goldstein R, Kunik ME, Yohannes AM, ACCP Workshop Panel on Anxiety and Depression in COPD. Anxiety and depression in COPD: current understanding, unanswered questions, and research needs. Chest 2008 Oct;134(4 Suppl):43S-56S [FREE Full text] [doi: 10.1378/chest.08-0342] [Medline: 18842932]
16. Lecheler L, Richter M, Franzen DP, Rampini SK, Cheetham M, Jenewein J, et al. The frequent and underrecognised co-occurrence of acute exacerbated COPD and depression warrants screening: a systematic review. Eur Respir Rev 2017 Jun 30;26(144) [FREE Full text] [doi: 10.1183/16000617.0026-2017] [Medline: 28615309]
17. Hill K, Geist R, Goldstein RS, Lacasse Y. Anxiety and depression in end-stage COPD. Eur Respir J 2008 Mar;31(3):667-677 [FREE Full text] [doi: 10.1183/09031936.00125707] [Medline: 18310400]
18. Dalal AA, Shah M, Lunacsek O, Hanania NA. Clinical and economic burden of depression/anxiety in chronic obstructive pulmonary disease patients within a managed care population. COPD 2011 Aug;8(4):293-299. [doi: 10.3109/15412555.2011.586659] [Medline: 21827298]
19. Quint JK, Baghai-Ravary R, Donaldson GC, Wedzicha JA. Relationship between depression and exacerbations in COPD. Eur Respir J 2008 Jul;32(1):53-60 [FREE Full text] [doi: 10.1183/09031936.00120107] [Medline: 18321938]
20. Gudmundsson G, Gislason T, Janson C, Lindberg E, Hallin R, Ulrik CS, et al. Risk factors for rehospitalisation in COPD: role of health status, anxiety and depression. Eur Respir J 2005 Sep;26(3):414-419 [FREE Full text] [doi: 10.1183/09031936.05.00078504] [Medline: 16135721]
21. Quanjer PH, Lebowitz MD, Gregg I, Miller MR, Pedersen OF. Peak expiratory flow: conclusions and recommendations of a Working Party of the European Respiratory Society. Eur Respir J Suppl 1997 Feb;24:2S-8S. [Medline: 9098701]
22. Jackson H, Hubbard R. Detecting chronic obstructive pulmonary disease using peak flow rate: cross sectional survey. BMJ 2003 Sep 20;327(7416):653-654 [FREE Full text] [doi: 10.1136/bmj.327.7416.653] [Medline: 14500437]
23. Casale R, Pasqualetti P. Cosinor analysis of circadian peak expiratory flow variability in normal subjects, passive smokers, heavy smokers, patients with chronic obstructive pulmonary disease and patients with interstitial lung disease. Respiration 1997;64(4):251-256. [doi: 10.1159/000196682] [Medline: 9257358]
24. Jones PW, Harding G, Berry P, Wiklund I, Chen W, Kline Leidy N. Development and first validation of the COPD Assessment Test. Eur Respir J 2009 Sep;34(3):648-654 [FREE Full text] [doi: 10.1183/09031936.00102509] [Medline: 19720809]
25. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
26. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]

27. Wang Z, Huang H, Cui L, Chen J, An J, Duan H, et al. Using Natural Language Processing Techniques to Provide Personalized Educational Materials for Chronic Disease Patients in China: Development and Assessment of a Knowledge-Based Health Recommender System. JMIR Med Inform 2020 Apr 23;8(4):e17642 [FREE Full text] [doi: 10.2196/17642] [Medline: 32324148]
28. Aaron SD, Donaldson GC, Whitmore GA, Hurst JR, Ramsay T, Wedzicha JA. Time course and pattern of COPD exacerbation onset. Thorax 2011 Oct 18;67(3):238-243. [doi: 10.1136/thoraxjnl-2011-200768] [Medline: 22008189]
29. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
30. Chen S, Chiu H, Xu B, Ma Y, Jin T, Wu M, et al. Reliability and validity of the PHQ-9 for screening late-life depression in Chinese primary care. Int J Geriatr Psychiatry 2010 Nov;25(11):1127-1133. [doi: 10.1002/gps.2442] [Medline: 20029795]
31. Lim S, Lam DC, Muttalif AR, Yunus F, Wongtim S, Lan LTT, et al. Impact of chronic obstructive pulmonary disease (COPD) in the Asia-Pacific region: the EPIC Asia population-based survey. Asia Pac Fam Med 2015;14(1):4 [FREE Full text] [doi: 10.1186/s12930-015-0020-9] [Medline: 25937817]
32. Tabak M, Brusse-Keizer M, van der Valk P, Hermens H, Vollenbroek-Hutten M. A telehealth program for self-management of COPD exacerbations and promotion of an active lifestyle: a pilot randomized controlled trial. Int J Chron Obstruct Pulmon Dis 2014;9:935-944 [FREE Full text] [doi: 10.2147/COPD.S60179] [Medline: 25246781]
33. Segrelles Calvo G, Gómez-Suárez C, Soriano JB, Zamora E, Gónzalez-Gamarra A, González-Béjar M, et al. A home telehealth program for patients with severe COPD: the PROMETE study. Respir Med 2014 Mar;108(3):453-462 [FREE Full text] [doi: 10.1016/j.rmed.2013.12.003] [Medline: 24433744]
34. Rixon L, Hirani SP, Cartwright M, Beynon M, Doll H, Steventon A, et al. A RCT of telehealth for COPD patient's quality of life: the whole system demonstrator evaluation. Clin Respir J 2017 Jul;11(4):459-469. [doi: 10.1111/crj.12359] [Medline: 26260325]
35. Jakobsen AS, Laursen LC, Rydahl-Hansen S, Østergaard B, Gerds TA, Emme C, et al. Home-based telehealth hospitalization for exacerbation of chronic obstructive pulmonary disease: findings from "the virtual hospital" trial. Telemed J E Health 2015 May;21(5):364-373 [FREE Full text] [doi: 10.1089/tmj.2014.0098] [Medline: 25654366]
36. 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]
37. Guo Y, Zhang X, Chen J. Value of CAT and mMRC in clinical evaluation of chronic obstructive pulmonary disease. Chongqing Medicine 2014;43(11):1304-1307 [FREE Full text] [doi: 10.1093/med/9780199358274.003.0033]
38. Dinesen B, Haesum LKE, Soerensen N, Nielsen C, Grann O, Hejlesen O, et al. Using preventive home monitoring to reduce hospital admission rates and reduce costs: a case study of telehealth among chronic obstructive pulmonary disease patients. J Telemed Telecare 2012 Jun;18(4):221-225. [doi: 10.1258/jtt.2012.110704] [Medline: 22653618]
39. Duan H, Wang Z, Ji Y, Ma L, Liu F, Chi M, et al. Using Goal-Directed Design to Create a Mobile Health App to Improve Patient Compliance With Hypertension Self-Management: Development and Deployment. JMIR Mhealth Uhealth 2020 Feb 25;8(2):e14466 [FREE Full text] [doi: 10.2196/14466] [Medline: 32130161]

Abbreviations

CAT: COPD Assessment Test
COPD: chronic obstructive pulmonary disease
EQ-5D: EuroQol-5D
FEV1: forced expiratory volume in 1 second
FVC: forced vital capacity
GAD-7: Generalized Anxiety Disorder-7
GOLD: Global Initiative for Chronic Obstructive Lung Disease
HIS: hospital information system
mHealth: mobile health
PEF: peak expiratory flow
PFT: pulmonary function test
PHQ-9: Patient Health Questionnaire 9
QoL: quality of life

Edited by G Eysenbach; submitted 01.09.19; peer-reviewed by J Dery, S Farag, C Batcho, R Wu; comments to author 15.04.20; revised version received 08.06.20; accepted 11.11.20; published 25.11.20.

Please cite as:
Deng N, Chen J, Liu Y, Wei S, Sheng L, Lu R, Wang Z, Zhu J, An J, Wang B, Lin H, Wang X, Zhou Y, Duan H, Ran P
Using Mobile Health Technology to Deliver a Community-Based Closed-Loop Management System for Chronic Obstructive Pulmonary Disease Patients in Remote Areas of China: Development and Prospective Observational Study
JMIR Mhealth Uhealth 2020;8(11):e15978
URL: http://mhealth.jmir.org/2020/11/e15978/
doi:10.2196/15978
PMID:33237036

Original Paper

Smartphone-Based Monitoring of Parkinson Disease: Quasi-Experimental Study to Quantify Hand Tremor Severity and Medication Effectiveness

Elina Kuosmanen[1], MSc; Florian Wolling[2], MSc; Julio Vega[3], PhD; Valerii Kan[1], MSc; Yuuki Nishiyama[4], PhD; Simon Harper[3], PhD; Kristof Van Laerhoven[2], PhD; Simo Hosio[1], PhD; Denzil Ferreira[1], PhD

[1]University of Oulu, Oulu, Finland

[2]University of Siegen, Siegen, Germany

[3]University of Manchester, Manchester, United Kingdom

[4]University of Tokyo, Tokyo, Japan

Corresponding Author:
Elina Kuosmanen, MSc
University of Oulu
Erkki Koiso-Kanttilan katu 3
P O Box 4500
Oulu, FI-90014
Finland
Phone: 358 504821517
Fax: 358 8 553 2612
Email: elina.kuosmanen@oulu.fi

Abstract

Background: Hand tremor typically has a negative impact on a person's ability to complete many common daily activities. Previous research has investigated how to quantify hand tremor with smartphones and wearable sensors, mainly under controlled data collection conditions. Solutions for daily real-life settings remain largely underexplored.

Objective: Our objective was to monitor and assess hand tremor severity in patients with Parkinson disease (PD), and to better understand the effects of PD medications in a naturalistic environment.

Methods: Using the Welch method, we generated periodograms of accelerometer data and computed signal features to compare patients with varying degrees of PD symptoms.

Results: We introduced and empirically evaluated the tremor intensity parameter (TIP), an accelerometer-based metric to quantify hand tremor severity in PD using smartphones. There was a statistically significant correlation between the TIP and self-assessed Unified Parkinson Disease Rating Scale (UPDRS) II tremor scores (Kendall rank correlation test: z=30.521, $P<.001$, $\tau=0.5367379$; n=11). An analysis of the "before" and "after" medication intake conditions identified a significant difference in accelerometer signal characteristics among participants with different levels of rigidity and bradykinesia (Wilcoxon rank sum test, $P<.05$).

Conclusions: Our work demonstrates the potential use of smartphone inertial sensors as a systematic symptom severity assessment mechanism to monitor PD symptoms and to assess medication effectiveness remotely. Our smartphone-based monitoring app may also be relevant for other conditions where hand tremor is a prevalent symptom.

(JMIR Mhealth Uhealth 2020;8(11):e21543) doi:10.2196/21543

KEYWORDS

Parkinson disease; smartphone; hand tremor; mobile health

Introduction

Background

Parkinson disease (PD) is a neurodegenerative condition that affects patients' physical and mental health [1,2] and has a wide variety of motor and nonmotor symptoms. Tremor is a cardinal motor symptom that can heavily hinder patients' quality of life [3] and is generally defined as an involuntary, rhythmic, oscillatory movement of a body part [4]. Tremor can be categorized based on its activation conditions into rest and action tremor; in turn, action tremor is further divided into kinetic, postural, or isometric subtypes [4]. Rest tremor affects body parts that are not being voluntarily activated [4], kinetic tremor appears during any voluntary movement, postural tremor presents while maintaining a posture against gravity [5], and isometric tremor occurs during a muscle contraction against a rigid surface [6].

Among patients with PD, approximately 75% suffer from rest tremor, around 50% from moderately severe postural tremor [7], and an undetermined percentage from kinetic tremor [6]. These three types of tremor are pivotal in understanding PD. Typically, the amplitude of rest tremor decreases when patients activate the affected muscles and increases during mentally stressful situations [4,8]. We target hand rest, postural, and kinetic tremor, which occur at different frequency ranges (3-6 Hz, 6-9 Hz, and 9-12 Hz, respectively) [5]. The severity of PD tremor is usually assessed visually by a health professional during clinical visits, using tools such as the Unified Parkinson Disease Rating Scale (UPDRS) [9]. Recently, however, researchers have investigated the use of unobtrusive and objective sensing technologies to detect and quantify hand tremor.

Dyskinesia is defined as involuntary movement, different from tremor, and is related to the timing and dosage of levodopa medication [10]. We refer to a movement in the 1-3 Hz frequency range as dyskinesia [11].

Related Work

Accelerometer data have been used to assess hand tremor using smartphones [11-17] and wearable devices [5,18-20]. Previous studies have attached an iPhone (Apple Inc) to a glove and collected data using the smartphone's built-in accelerometer [11,12,15]. In a study by LeMoyne et al [11], subjects were asked to use the glove while extending their forearm for 10 seconds. The study found a statistically significant difference in the frequency response of the acceleration signal between a participant with PD and one without. Barrantes et al [12] measured rest and postural tremor in 30-second episodes and identified relevant accelerometer features to classify PD tremor and essential tremor. Similarly, Duque et al [15] collected accelerometer data with participants at rest and with their arms extended, and utilized machine learning to classify PD and essential tremor.

Bazgir et al [16] used a similar setup with a glove to classify rest, postural, and kinetic tremor UPDRS scores based on accelerometer data logged during three scripted 1-minute tests: at rest, with arms stretched, and while touching their nose with their index finger. Kostikis et al [14] used a glove-mounted smartphone to quantify rest and postural tremor severity. They found a statistically significant difference between healthy participants and participants with PD, but not between the left and right hands of people with PD. In addition, they studied the effects of PD medications in two volunteers with PD using accelerometer data collected in "on" and "off" medication states in laboratory conditions. The measurements were taken for the right and left hands separately, at rest, and with their hands extended. The "off" measurement was taken right before medication intake, and the "on" measurement was taken 1 hour after intake. The researchers detected a decrease in the metrics (the sums of the squared magnitudes of the acceleration and the sum of absolute differences in the acceleration vector) after medication intake, with the exception of the right-hand extended task for one of the volunteers.

Although these previous experiments have had positive results on quantifying tremor, we believe that the utility of the findings outside of the laboratory or a health care facility is limited. The practicality of carrying and wearing a glove at all times is up for debate, especially under extreme weather conditions. Accordingly, a smartphone-only solution was first investigated by Woods et al [13], where subjects held a smartphone for 10 seconds with their arm perpendicular to their body and elbow pointing out under six conditions: with eyes open, with eyes closed, during a bubble-balancing task, during a laser-pointing task with two different distances, and while counting backward by decrements of three. The authors detected a statistically significant difference in the accelerometer signal between a group with PD and a group with essential tremor, but they focused on the effects of the six cognitive tasks on tremor. Following up this line of research, this paper explores the difference in the accelerometer signal captured during a scripted task without additional hardware other than a smartphone.

In our study, the data were collected in naturalistic settings using a mobile toolkit, the Sentient Tracking of Parkinson's (STOP) app, for monitoring PD symptoms in daily life. STOP includes a gamified tremor assessment module based on a ball-balancing game that logs the smartphone's accelerometer, gyroscope, and rotation data. STOP also provides users with a medication intake journal and a daily symptom survey mechanism [21-23]. The data set used in this paper was published previously [21]. In this study, we analyzed the data set's accelerometer and medication intake data to answer the following research questions: (1) how feasible is it to characterize hand tremor using inertial data captured during our smartphone game?, and (2) can the effects of PD medication be detected using the same inertial data captured during game sessions played before and after medication intake?

Methods

We installed the STOP app into the smartphones of 13 participants diagnosed with PD and collected accelerometer data and medication logs. We used the Welch method to generate the power spectral densities (PSDs) and extracted features from the accelerometer data that we used to investigate the feasibility of hand tremor assessment and medication effect.

STOP Application and Data Collection

STOP is a smartphone app developed for people with PD with four core functionalities: (1) an accelerometer-based ball game for quantifying patients' hand tremor, (2) a medication journal for logging medication intake times, (3) a daily survey for reporting the overall severity of PD symptoms, and (4) reminder notifications [21-23].

To play the ball game, one has to place the smartphone horizontally on the palm of the hand for 10 seconds and try to keep a virtual ball inside a circle at the center of the screen. During the game session, STOP logs data from the accelerometer, linear accelerometer (acceleration without gravity's influence), gyroscope, and rotation vector sensors [24]. It also records the position of the ball in relation to the inner circle's center and the screen's pixel density to compute an adjusted distance between the center of the ball and the center of the screen. The inertial sensors' sampling frequency is set to 50 Hz (or the device's maximum, if less than 50 Hz). In addition, users can record their medication intake using the "now" or "specify time" buttons, or with their voice via the natural language processing provided by Wit.ai [25].

During a real-world trial of STOP, data were collected from 13 participants with PD, eight females and five males [21]. The participants were recruited from two countries, seven from Finland and six from the United Kingdom. In this study, we had to exclude two participants because of poor data quality (see Data Set). Table 1 provides a summary of the remaining 11 participants' characteristics; more details are provided in Multimedia Appendix 1. Participants were asked to install and use the STOP app for 1 month on their personal smartphones (five iPhones and six Android phones) and to participate in three interviews (at the start of the study, midway, and at the end).

Participants from Finland were located around the country, and their consent to participate in the study was given via the application. Participants from the United Kingdom, on the other hand, signed a paper consent form. Following local guidelines, approval from the University of Oulu's ethical committee was not needed because the risks associated with participating in the study were similar to those of daily smartphone use. In previous publications, we have shared users' experience during the trial and an analysis of the interview data [21]. To summarize, participants were willing to use digital tools to track their condition and were open to the possibility of sharing their data with their doctors, which functioned as a motivator to use such tools. In this paper, we analyze the inertial sensor data recorded during the game sessions and medication logs to quantify the severity of hand tremor and the effect that medication has on it.

Table 1. Overview of participants' characteristics.

Characteristics	Participants (n=11)
Age (years), mean (range)	64.7 (52-73)
Years since PD diagnosed, mean (range)	7.1 (2-17)
Number of PD medications, mean (range)	3 (1-5)
Number of total daily medications[a], mean (range)	4.3 (1-7)
UPDRS II score[b], mean (range)	11.8 (3-31)
Tremor on UPDRS, mean (range)	1.2 (0-3)
Deep brain stimulator installed, n (%)	2 (18%)
Suffer from hand tremor, n (%)	4 (36%)
Plays with tremor-affected hand, n (%)	2 (18%)
Suffers from other issues affecting hands, n (%)	
Rigidity	3 (27%)
Bradykinesia	1 (9%)

[a]Refers to the number of medication intake times (ie, how many times per day the participant has to take medications, one or several at a time).

[b]The Unified Parkinson Disease Rating Scale (UPDRS) II score quantifies the severity level of Parkinson disease (PD) symptoms affecting daily activities (maximum score of 52). The scale for the tremor item on the UPDRS is as follows: 0=no tremor, 1=slight and infrequently present tremor, 2=moderate and bothersome tremor, 3=severe tremor interfering with many activities, and 4=marked tremor interfering with most activities.

Data Set

Our data set contained a total of 1856 medication logs (mean of 107 [SD 54.9] logs per participant) and 2213 game sessions (mean of 138 [SD 60.6] sessions per participant). These data were recorded in 13 participants (P01 to P13) in naturalistic conditions. Participants had varying medication regimens.

Game sessions were 10 seconds long. We excluded P03's sessions because the accelerometer sampling rate of his smartphone was approximately 25 Hz instead of the desired 50 Hz, and we excluded P04's sessions because they only contained one sensor sample throughout the entire game for unknown technical reasons. P05's phone had data synchronization issues, so only 1 week of data was collected, and P07 missed the first week of data collection because he had problems installing the application. Despite this, we included P05 and P07 in the analysis, resulting in a total of 11 participants.

Because the data were collected during a trial deployment of the STOP app, there were no participant exclusion criteria related to PD symptom severity. Based on the UPDRS II tremor self-reports, we categorized the participants into five groups: (1) all participants: P01, P02, P05, P06, P07, P08, P09, P10, P11, P12, and P13; (2) no tremor (participants reported no tremor): P02, P09, and P11; (3) tremor (participants reported tremor but in an unspecified location): P06, P07, P08, and P12; (4) hand tremor (participants reported hand tremor and played with the unaffected hand): P01 and P05; and (5) plays with hand tremor (participants reported hand tremor and played with the affected hand): P10 and P13.

We highlighted individual circumstances that might affect STOP's measurements. P02 reported that his hand rigidity helped him to keep the ball still during game sessions. P09 had poor rotation in his wrists and P11 suffered from rigidity and bradykinesia that make him feel stiff and slow, which might have had a similar effect to that of P02. Finally, P01 was right-handed but used his left hand for playing. Multimedia Appendix 1 provides more details about participants' symptoms and playing conditions.

Data Preprocessing

Accelerometer data were recorded as participants played a game for a duration of 10 seconds, henceforth referred to as a "game session." The accelerometer sampling rate was set to 50 Hz, but the sampling rate varied across different smartphones, as the participants used their own devices for the study. In addition, for some devices, the sampling rate varied within a game session. In the examples in Figure 1, P01's sampling rate stayed close to the requested rate, while the rate varied in P08's device. For all devices, the sampling frequency was set to 50 Hz. To address the variation in the sensors' sampling rate and get uniformly sampled data, we applied a linear interpolation on the accelerometer signal (see Multimedia Appendix 2 for technical details).

We identified the closest medication intake record to each game session and labeled the game sessions as "before" or "after" (see Figure 2). Depending on multiple factors, medication can take at least 15 minutes to kick in [26]. Therefore, for participants who had to take their medication once or twice per day (every 24 or 12 hours), a game session was considered "before" medication if it was played between 5 hours before or 15 minutes after the intake log was completed. In contrast, the game session was labeled "after" medication if it was played between 30 minutes and 3 hours after the medication intake log was completed. For participants with more than two medication intakes per day, we used shorter thresholds. Specifically, the game session was labeled "before" if it was played either 1 hour before or 15 minutes after the medication intake, and "after" if it was played between 30 and 90 minutes after the medication intake. Game sessions outside of these periods were not included in our medication effect analysis. Figure 2 shows an overview of the labeled time periods; the medication intake time is denoted as a red solid vertical line centered around zero.

Figure 1. Best case (left) and worst case (right) examples of varying smartphone sampling frequency during a game session.

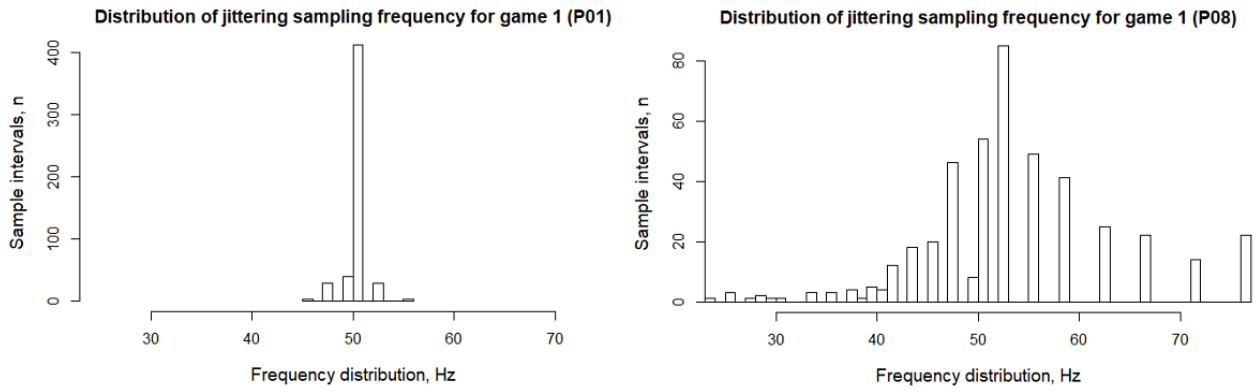
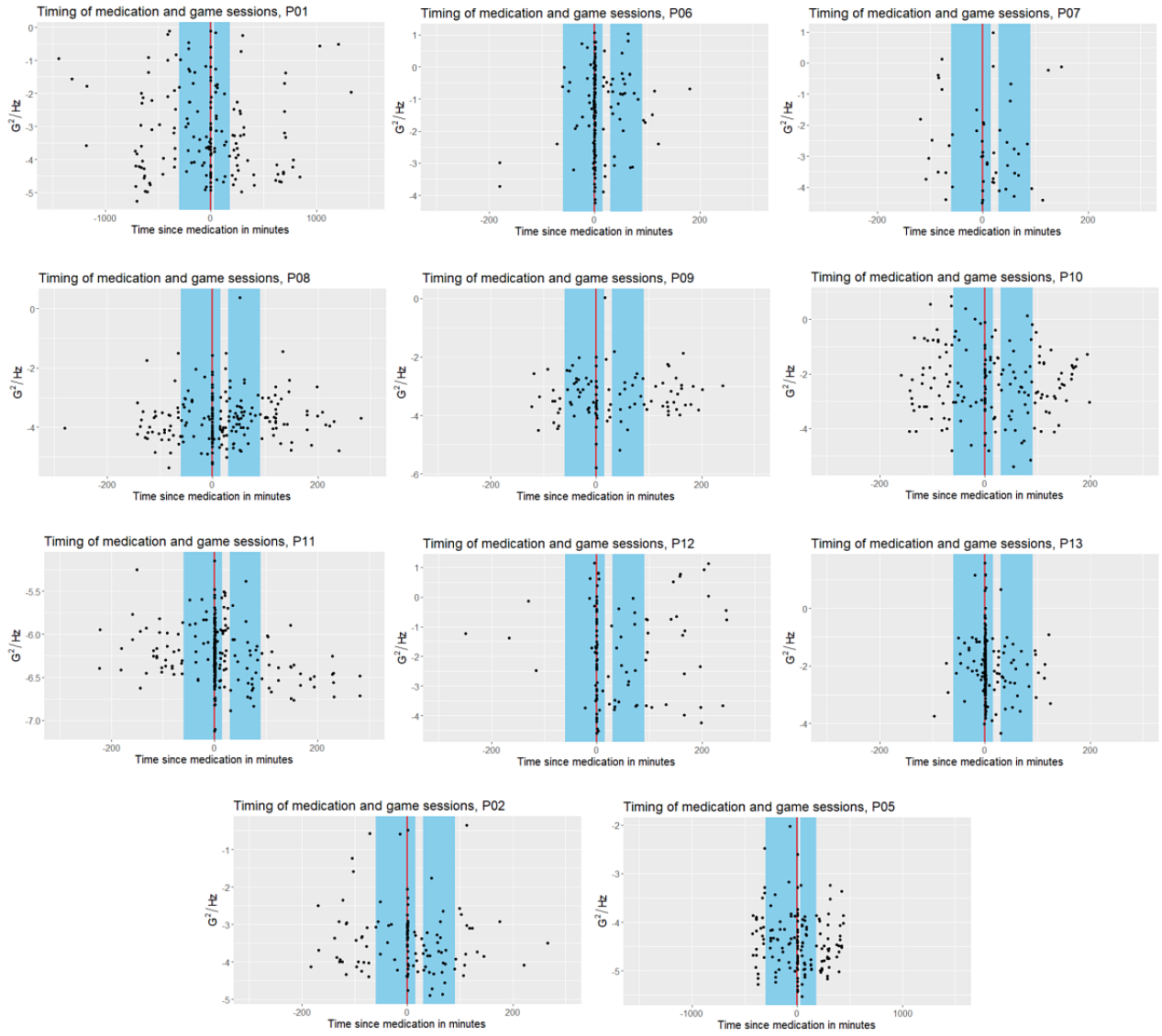


Figure 2. The timing of medication intake and game sessions. The x-axis shows the time since medication, 0 is the medication intake time and is highlighted with a red vertical line. Each game is associated with the closest medication intake time, either before or after. The y-axis presents the acceleration signal power in logarithmic scale; the sum of power is calculated over the entire spectrum for each game session. Note that the y-axis ranges differ. The first three rows show participants with more than two intakes per day while the last row shows those with only one or two.



Frequency Analysis: PSD

PD symptoms can be observed in specific frequency bands: dyskinesia (1-3 Hz), rest tremor (3-6 Hz), postural tremor (6-9 Hz), and kinetic tremor (9-12 Hz). As described in the introduction, tremor can be classified by its activation condition. In our study setup, depending on the user's posture during a game session, we expected to see differences in the accelerometer signal in rest tremor, postural tremor, and dyskinesia frequencies, which we tried to detect by analyzing this signal in the frequency domain.

We used the Welch method [27] to generate periodograms of every participant's game sequences. This method generates a nonparametric estimation of the PSD, determining the power contained in the signal's particular frequency components (see Multimedia Appendix 2 for more details). The left columns in Figures 3-5 show the mean of all games' periodograms as well as the confidence interval around the mean. The right columns depict the averaged periodograms for the "before" and "after" game subsets, respectively. We observed a higher PSD value in the groups with PD tremor (ie, "tremor," "hand tremor," and "plays with hand tremor" groups) than in the "no tremor" group. A comparison of the groups is presented in the results section.

From the periodograms, several features were calculated to describe the characteristics of the signal:

- area under the curve (AUC): describes the total power of the signal (Hz) [12] (Figures 3-5 present the mean of PSDs in each frequency);
- peak value (PV): represents the maximum value of the PSD;
- fundamental frequency (F0): the frequency of maximum power [5,12,16,19]. The F0 can be used to categorize the game sessions as dyskinesia, rest tremor, postural tremor, or kinetic tremor games [5,19]. The percentage of game sessions in each category of each participant is summarized in Table 2 (the red line in Figure 6 illustrates the F0);

- central frequency (F50): the central point where the periodogram is divided into two equal parts in PSD [5,12,16,19] (the green line in Figure 6 illustrates the F50);
- frequency dispersion (SF50): describes the width of the frequency band around F50 containing 68% of the total power of the signal [5,16,19] (see the blue area in Figure 6);
- |F50-F0|: the difference between F50 and F0 [5,16,19] (see the distance between the red [F0] and green [F50] lines in Figure 6);
- tremor intensity parameter (TIP): calculated as PV divided by SF50. In Figure 6, P10 has a narrow, high peak in PSD, causing a high TIP, whereas P02 has a lower PV and a wide SF50, causing a low TIP. We introduce this parameter to quantify tremor severity based on accelerometer data—a higher TIP indicates a more severe tremor.

We utilized these features to quantify tremor severity during a game session and to detect a difference in medication effects between different game sessions.

Figure 3. Mean of the power spectral densities with the 95% CI for the "no tremor" group (P02, P09, and P11). The left column shows the mean of all game sessions, and the right column shows the mean of the power spectral densities for "before" (red) and "after" (blue) games. Note that the y-axis ranges differ. Frequency areas (dyskinesia, rest tremor, postural tremor, and kinetic tremor) are denoted by different column shades of the background.

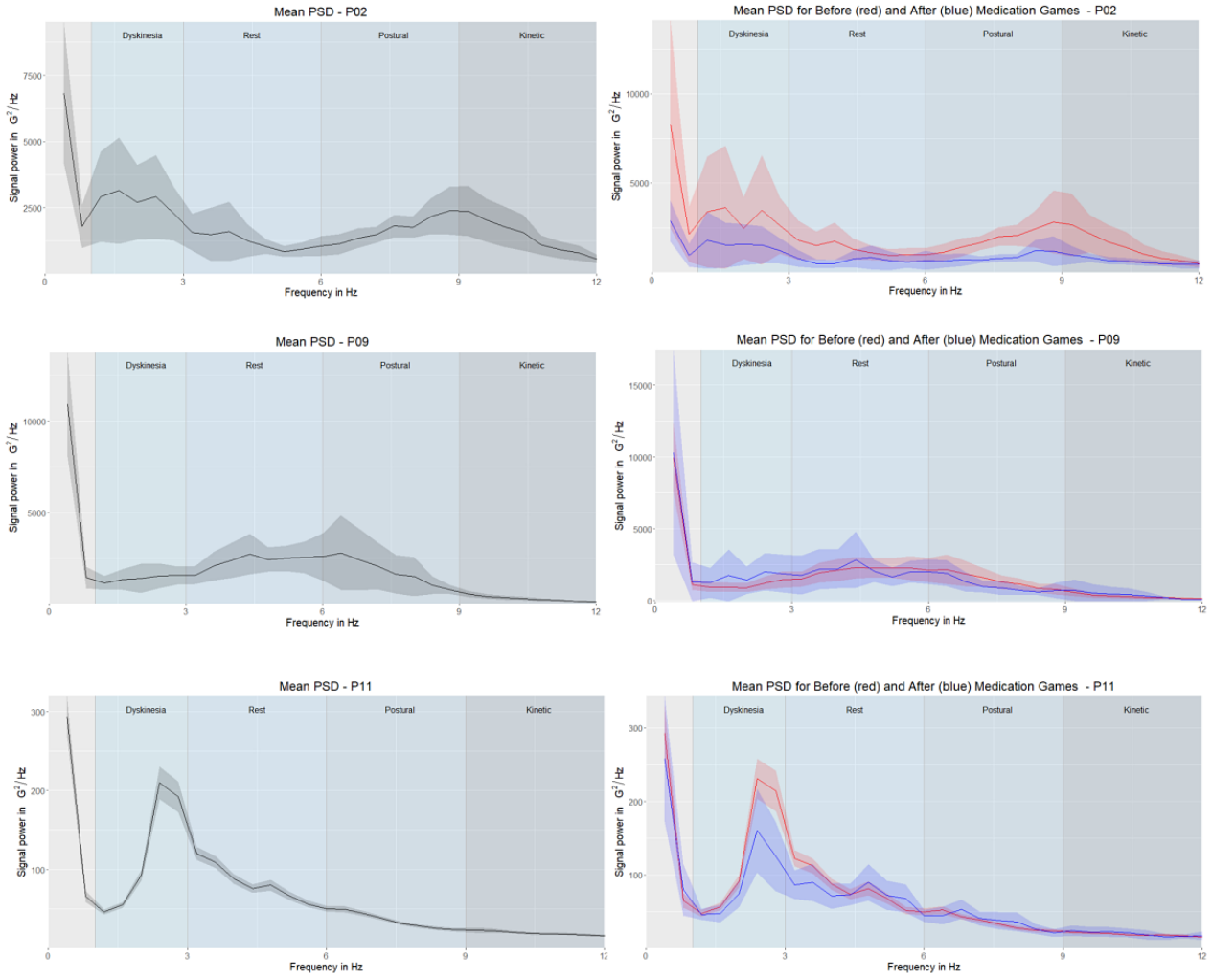


Figure 4. Mean of the power spectral densities with the 95% CI for the "tremor" group (P06, P07, P08, and P12). The left column shows the mean of all game sessions, and the right column shows the mean of the power spectral densities for "before" (red) and "after" (blue) games. Note that the y-axis ranges differ. Frequency areas (dyskinesia, rest tremor, postural tremor, and kinetic tremor) are denoted by different column shades of the background.

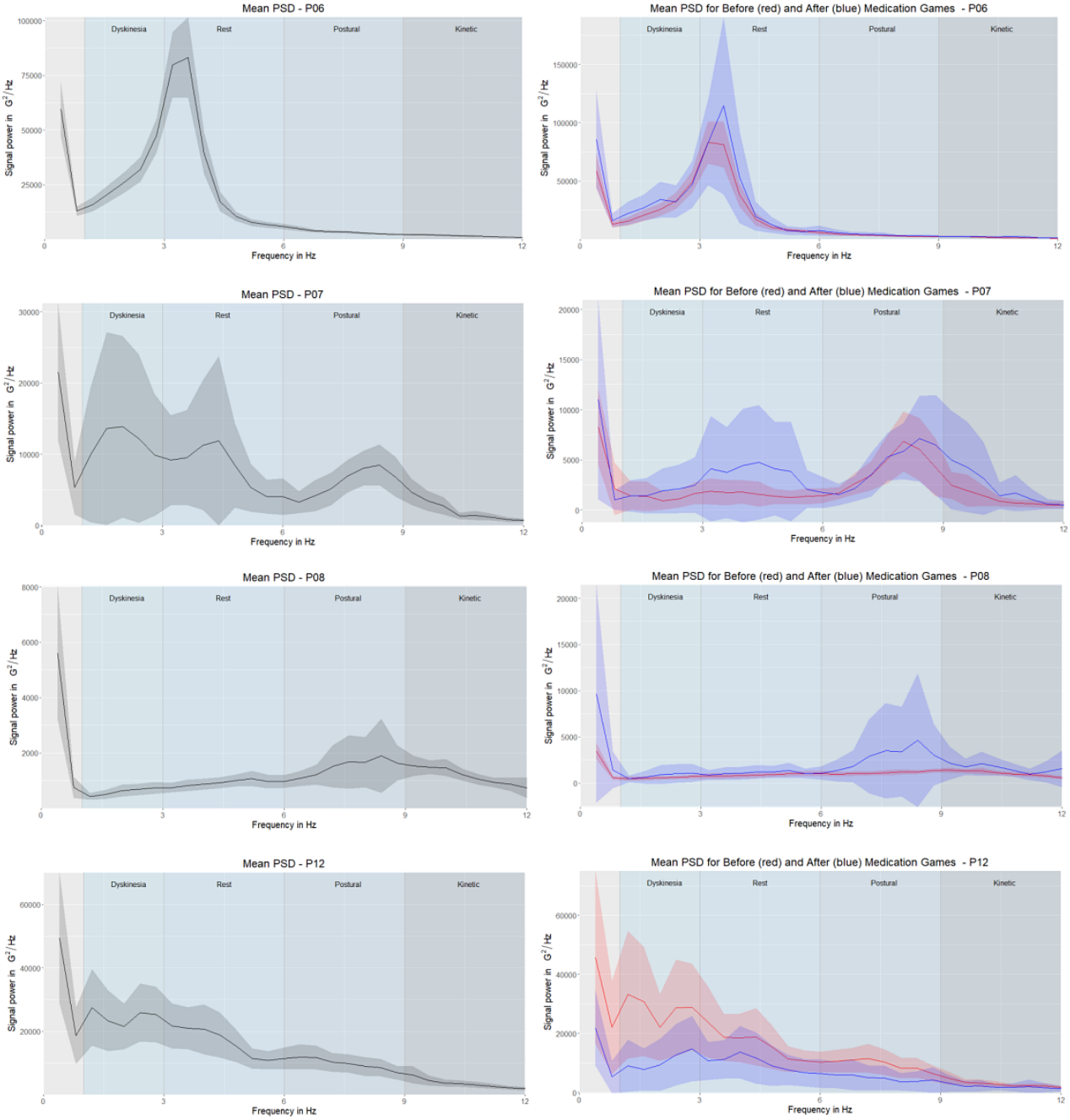


Figure 5. Mean of the power spectral densities (PSDs) with the 95% CI for the "hand tremor" group (P01 and P05) and the "plays with hand tremor" group (P10 and P13). The left column shows the mean of all game sessions, and the right column shows the mean of the power spectral densities for "before" (red) and "after" (blue) games. Note that the y-axis ranges differ. Frequency areas (dyskinesia, rest tremor, postural tremor, and kinetic tremor) are denoted by different column shades of the background.

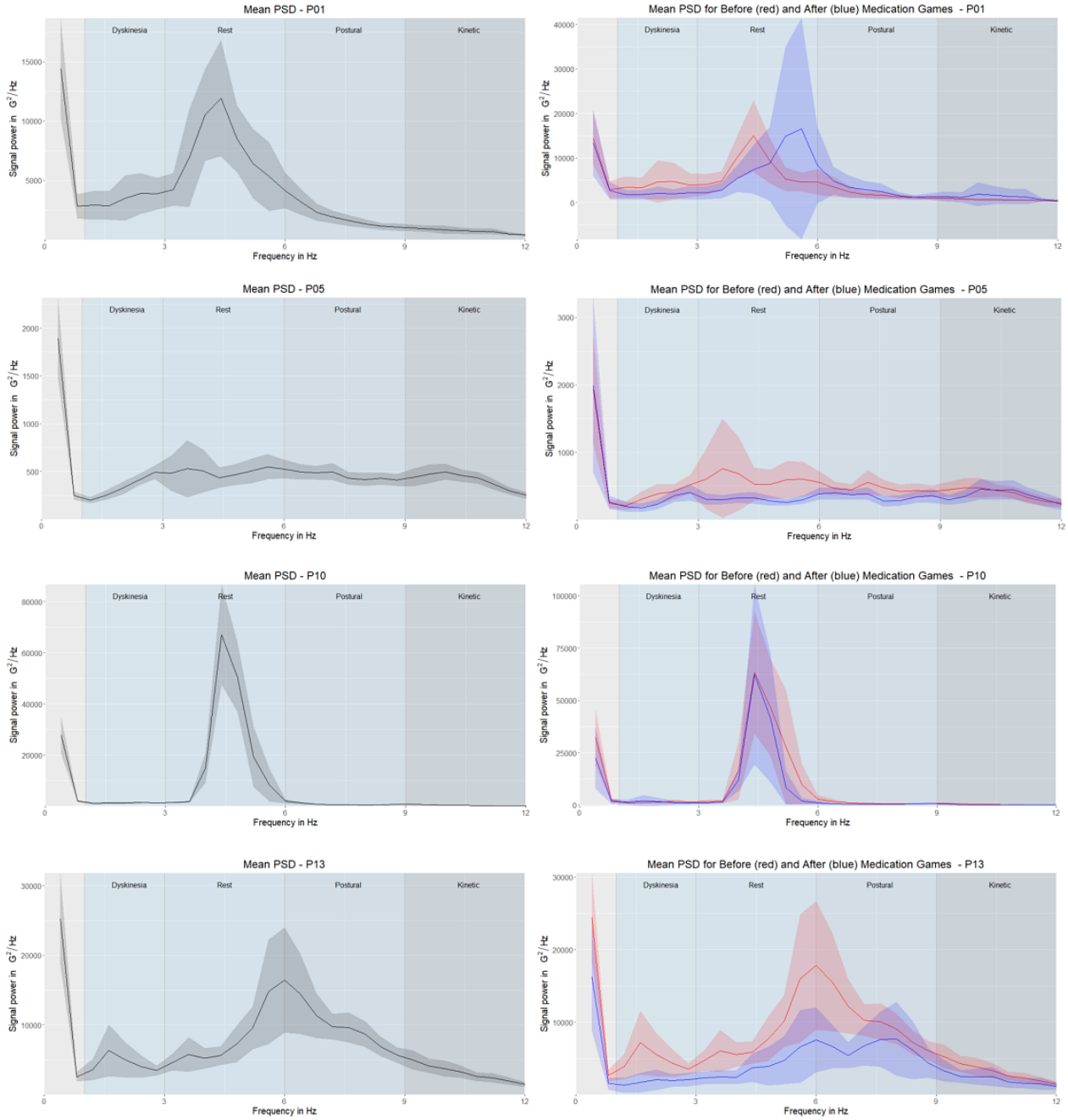
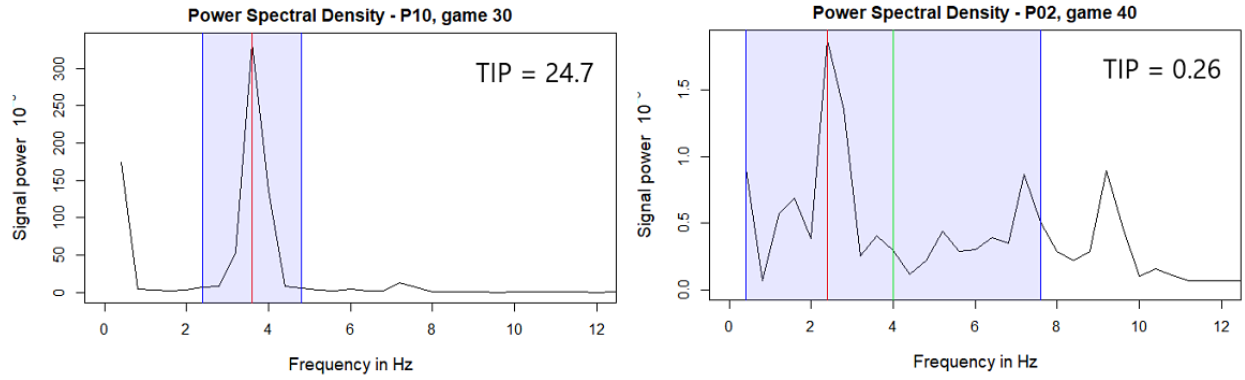


Table 2. The game sessions categorized as dyskinesia, rest tremor, postural tremor, or kinetic tremor according to the fundamental frequency are shown as percentages of all game sessions of each participant (the absolute number of game sessions appears in parentheses).

Participant by group	Dyskinesia	Rest tremor	Postural tremor	Kinetic tremor
No tremor				
P02	52%[a] (56/107)	7% (7/107)	41% (44/107)	0% (0/107)
P09	18% (19/104)	62%[a] (64/104)	20% (21/104)	0% (0/104)
P11	64%[a] (170/265)	34% (89/265)	2% (5/265)	0% (1/265)
Tremor				
P06	30% (53/175)	70%[a] (122/175)	0% (0/175)	0% (0/175)
P07	10% (5/51)	10% (5/51)	80%[a] (41/51)	0% (0/51)
P08	30% (52/174)	12% (21/174)	58%[a] (101/174)	0% (0/174)
P12	34% (38/111)	40%[a] (44/111)	26% (29/111)	0% (0/111)
Hand tremor				
P01	1% (26/167)	66%[a] (111/167)	18% (30/167)	0% (0/167)
P05	41%[a] (63/152)	21% (32/152)	38% (57/152)	0% (0/152)
Plays with hand tremor				
P10	4% (7/169)	95%[a] (160/169)	1% (2/169)	0% (0/169)
P13	8% (22/282)	19% (54/282)	73%[a] (206/282)	0% (0/282)

[a]The most prevalent symptom of the participant.

Figure 6. The power spectral density of one game of P10, playing with hand tremor, and one game of P02, with no tremor. The red, vertical line shows the fundamental frequency (F0), the green line shows the central frequency (F50), and the gap between the lines is the difference between F50 and F0 (|F50-F0|). For P10, F0 and F50 are the same frequency, hence, |F50-F0|=0. The blue rectangle shows the SF50 (the frequency band around F50 containing 68% of the total power of the signal). P10 has a high peak value (PV) and a narrow SF50, leading to a high tremor intensity parameter (TIP) of 24.7. The PV of P02 is small (as is the signal power in the PSD in general) and SF50 is wide; hence, he has a low TIP of 0.26. Note that the y-axis ranges in both plots differ.



Results

In this section, we study our two research questions using the PSD features described in the previous section: (1) how feasible is it to characterize tremor using inertial data captured during our smartphone game?, and (2) can the effects of PD medication be detected using the same inertial data captured during game sessions played before and after medication?

Hand Tremor Characterization via the TIP: Proposal for an Objective Hand Tremor Severity Score

We found a significant correlation between self-reported UPDRS II tremor severity scores (0 to 4) and the TIP (Kendall rank correlation test: z=30.521, P<.001, τ=0.5367379; n=11). UPDRS II tremor scores and descriptive statistics of the TIP for each participant are shown in Table 3.

We then compared the groups across all features (see Figures 7 and 8). Figure 7 shows the AUC for all frequency ranges. We found that for dyskinesia the means of all four groups were similar and the tremor group had the largest variability. For the

other three frequency ranges (rest tremor, postural tremor, and kinetic tremor), the mean of the "plays with hand tremor" group was greater than that of the other groups. The participants in the "plays with hand tremor" group also had the highest PV and lowest SF50, and thus the highest TIP score (Figure 8). We used a Wilcoxon rank sum test to confirm that the differences between our 4 groups were statistically significant, resulting in six pairwise comparisons for AUC for all four frequency areas (dyskinesia, rest tremor, postural tremor, and kinetic tremor), PV, F0, F50, SF50, |F50-F0| and TIP (see *P* values in Table 4). Table 4 is extended in Multimedia Appendix 3, also providing the W for the Wilcoxon rank sum test.

All features were significantly different between the "no tremor" and "plays with hand tremor" groups. Additionally, all features except for SF50 showed a significant difference between the "no tremor" and "tremor" groups and between the "no tremor" and "hand tremor" groups. SF50 describes the width of the frequency band around F50 containing 68% of the total power of the signal. This suggests that when the tremor was located in a body part other than the hand holding the device, the power of the signal was spread in a wider frequency range, resembling the case with no tremor. However, the "no tremor" group differed significantly from the groups with tremors.

Features between the "tremor" and "hand tremor" groups were significantly different only in the AUC for the dyskinesia, postural tremor, and kinetic tremor frequency ranges. Hence, we can say that the effect of tremors on the accelerometer signal in these groups was mainly similar. In contrast, the comparison of the "plays with hand tremor" group with the "tremor" group and the "hand tremor" group showed significant differences in all features except F0 and F50. The tremor effect was similar in frequency, but the magnitude of the tremors was different when the tremor hand was used for playing.

Table 3. Self-reported tremor severity scores using the Unified Parkinson Disease Rating Scale (UPDRS) II, item 16.

		Distribution of tremor intensity parameter					
Participant	UPDRS II item 16 score[a]	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
P02	0	0.05	0.17	0.33	0.97	0.66	22.00
P09	0	0.04	0.33	0.56	1.09	1.18	19.10
P11	0	0.01	0.03	0.04	0.05	0.06	0.31
P05	1	0.03	0.07	0.10	0.18	0.16	5.02
P07	1	0.15	0.45	1.40	5.49	2.65	81.24
P08	1	0.04	0.12	0.21	0.39	0.38	14.57
P12	1	0.07	0.71	2.05	9.47	8.81	142.83
P01	2	0.03	0.20	0.84	7.41	4.04	177.01
P10	2	0.06	3.32	15.93	59.26	59.65	769.10
P13	2	0.15	0.92	1.69	6.40	3.79	610.81
P06	3	0.31	4.97	16.54	35.48	38.12	321.28

[a]0=no tremor, 1=slight and infrequently present tremor, 2=moderate and bothersome tremor, 3=severe tremor interfering with many activities, and 4=marked tremor interfering with most activities.

Figure 7. Comparison of tremor groups using area under the curve for each frequency range: dyskinesia (1-3 Hz), rest tremor (3-6 Hz), postural tremor (6-9 Hz), and kinetic tremor (9-12 Hz).

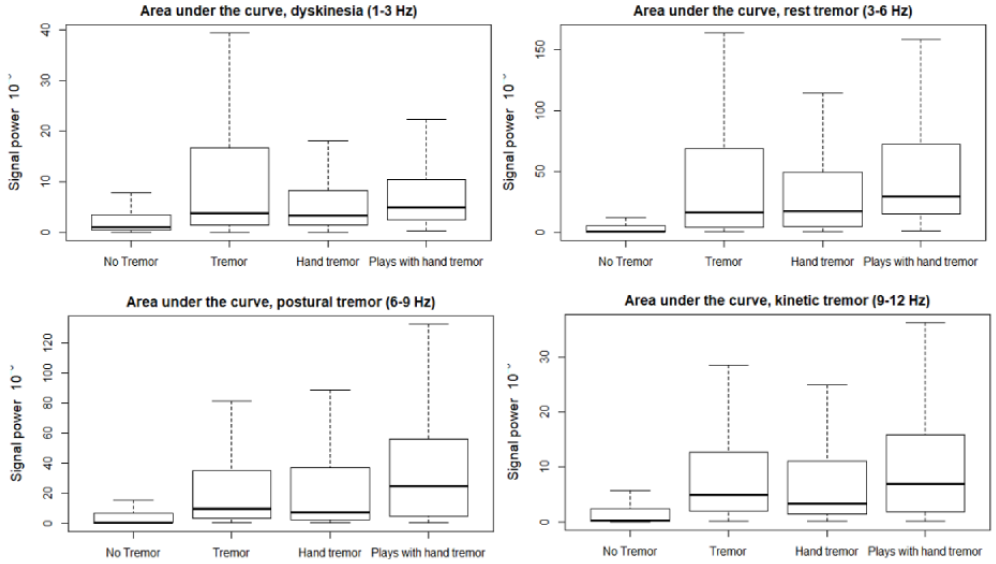


Figure 8. Comparison of groups in terms of the following features: peak value, fundamental frequency (F0), central frequency (F50), frequency dispersion (SF50), difference between F50 and F0 (|F50-F0|), and tremor intensity parameter.

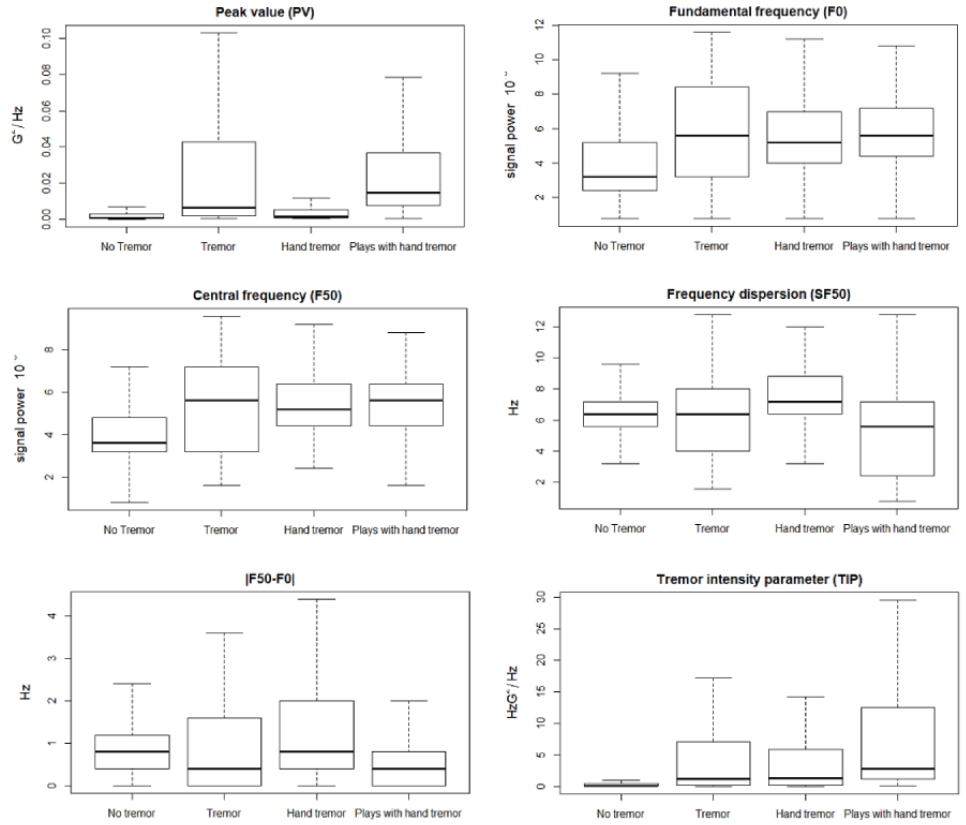


Table 4. The *P* values from Wilcoxon rank sum tests comparing groups for each feature: area under the curve (AUC) for all four frequency areas, peak value, fundamental frequency (F0), central frequency (F50), frequency dispersion (SF50), difference between F50 and F0 (|F50-F0|), and tremor intensity parameter (TIP).

	No tremor vs tremor	No tremor vs hand tremor	No tremor vs plays with hand tremor	Tremor vs hand tremor	Tremor vs plays with hand tremor	Hand tremor vs plays with hand tremor
AUC						
Dyskinesia	<.001	<.001	<.001	<.001	.03	<.001
Rest tremor	<.001	<.001	<.001	.89	<.001	<.001
Postural tremor	<.001	<.001	<.001	.02	<.001	<.001
Kinetic tremor	<.001	<.001	<.001	<.001	.03	<.001
Peak value	<.001	<.001	<.001	.49	<.001	<.001
F0	<.001	<.001	<.001	.62	.23	.30
F50	<.001	<.001	<.001	.43	.15	.35
SF50	.71	.54	<.001	.39	<.001	<.001
|F50 - F0|	<.001	<.001	<.001	.18	<.001	<.001
TIP	<.001	<.001	<.001	.78	<.001	<.001

Medication Effect Detection

We investigated the effect of medication intake on the accelerometer signal characteristics. PD medication is often targeted to alleviate motor symptoms; thus, it could have affected participants' motor performance during their game sessions. To explore this possibility, we compared the "before" and "after" game sessions of each individual.

In Figures 3-5, on the graphs on the right-hand side, we highlighted the mean PSD of "before" (red) and "after" (blue) game sessions with 95% CIs. Because our sample was relatively small, some of the 95% CI boundaries were negative [28]. In the "no tremor" group (Figure 3), P02 and P11 had peaks in the dyskinesia frequency range, and the mean PSD of "before" games was higher than the mean PSD of "after" games. This suggests that the medication partially alleviated this symptom. The 95% CIs for P09 were mostly overlapping, suggesting that there was no difference in the mean PSD of "before" and "after" games.

In the "tremor" group (Figure 4), P06 had a high peak in the rest tremor frequency area, P08 had a high peak in the postural tremor frequency area, and P07 had a peak in both. Compared with the "no tremor" group, the peaks in the "tremor" group were in tremor frequency ranges, rather than in the dyskinesia frequency range, which matches our expectation of observing this symptom. Even though the 95 CIs were overlapping, the 95% CI for "before" games was narrower. Figure 2 shows that P06 and P08 often played the game at the same time as medication intake, which might have resulted in the narrowing of the 95% CI for "before" games. P12 had two peaks in the "before" sessions in the dyskinesia frequency range; these peaks were lower in "after" sessions.

In the "hand tremor" and "plays with hand tremor" groups (Figure 5), P10 had a clear peak in the rest tremor frequency range. However, the effect of medication was not visible, since the "before" and "after" 95% CIs fully overlapped. This might indicate that the medication was working well, and its effect was maintained prior to the next intake. For P01, who suffered from hand tremor but played with his nonaffected hand, we found a difference in the tremor frequency between "before" and "after" game sessions, and the peak frequency had shifted (Figure 5). For P05 and P13, we observed that the mean PSD of "before" games was higher than of "after" games (Figure 5). P05 had a narrow 95% CI for "after" game sessions; thus, the performance was more predictable after medication intake.

In Table 5, we summarized the changes detected in all of the features between "before" and "after" game sessions. Many participants had changes in their features between the game sessions. A Wilcoxon rank sum test confirmed statistically significant differences for three participants: (1) P02 in AUC dyskinesia (W=861, *P*=.005), AUC resting tremor (W=1016, *P*<.001), AUC postural tremor (W=970, *P*=.002), AUC kinetic tremor (W=872, *P*=.036), PV (W=949, *P*=.004), SF50 (W=421, *P*=.006), and TIP (W=953, *P*=.003); (2) P09 in SF50 (W=469, *P*=.005); and (3) P11 in AUC dyskinesia (W=2767, *P*=.003), PV (W=2590, *P*=.024), F50 (W=1490, *P*=.027), SF50 (W=1327, *P*=.004), and TIP (W=2665, *P*=.011). These three participants (P02, P09, and P11) reported no tremor (UPDRS II, item 16). P02 and P09 presented with rigidity, and P11 presented with rigidity and bradykinesia; hence, it seems that the medication effect was more visible for these symptoms.

Table 5. Change in means of features as percentages between "before" and "after" medication game sessions. The negative values represent a lower mean in "after" game sessions than in "before" game sessions, while positive values represent the opposite.

Participant by group	AUC[a], dyskinesia	AUC, rest tremor	AUC, postural tremor	AUC, kinetic tremor	PV[b]	F0[c]	F50[d]	SF50[e]	|F50-F0|[f]	TIP[g]
No tremor										
P02	–51[h]	–51[i]	–53[h]	–56[h]	–53[h]	10	9	11[h]	24	–3[h]
P09	52	1	–23	30	1	–13	–14	–23[h]	–28	21
P11	–29[h]	–8	4	4	–25[h]	7	10[h]	10[h]	48	–35[h]
Tremor										
P06	14	21	21	25	22	3	–4	5	31	4
P07	45	148	8	102	4	10	2	1	2	–8
P08	41	28	142	45	139	1	0	3	4	108
P12	–62	–40	–48	–30	–59	–14	–9	0	–4	–69
Hand tremor										
P01	–52	9	51	104	2	7	9	–5	–24	–3
P05	–26	–52	–26	–7	–45	13	–3	1	23	–51
Plays with hand tremor										
P10	1	–23	–52	–42	–21	–1	–2	7	61	–3
P13	–63	–53	–41	–34	–51	–3	1	1	–27	–65

[a]AUC: area under the curve.

[b]PV: peak value.

[c]F0: fundamental frequency.

[d]F50: central frequency.

[e]SF50: frequency dispersion.

[f]|F50-F0|: difference between F50 and F0.

[g]TIP: tremor intensity parameter.

[h]Difference is statistically significant at $P<.05$, based on the Wilcoxon rank sum test.

[i]Difference is statistically significant at $P<.001$, based on the Wilcoxon rank sum test.

Discussion

In this paper, we show that it is feasible to detect and characterize PD hand tremor severity using accelerometer data captured during game play. Further, we investigated the medication effect on the accelerometer signal, demonstrating a statistically significant difference in the accelerometer data characteristics of the game sessions played before and after medication intake by participants with rigidity and bradykinesia.

Revisiting the Research Questions

First, how feasible is it to characterize hand tremor using inertial data captured during our smartphone game? To this end, we introduced the TIP for characterizing hand tremor severity, as computed using accelerometer data. We show that TIP is significantly correlated with the tremor score (item 16) on the UPDRS II [9]. TIP was significantly different between participants with no tremor and those with tremor symptoms, as well as between the participants playing with the tremor hand and participants with tremor in the opposite hand or in other body parts (Table 4). These results suggest that it is possible to objectively detect and quantify the severity of hand tremor using smartphone accelerometer data across different tremor types and intensities.

Inspired by previous work in hand tremor analysis using accelerometer data [5,11-16,18,19], we analyzed the accelerometer data collected during a 1-month real-world trial of the STOP app [21]. Earlier studies have already shown that accelerometer signals can be used to measure tremor under controlled conditions similar to traditional clinical assessments using the UPDRS II, either by discriminating between people with and without PD [11-13,15,18] or by measuring tremor severity [5,14,16,19]. We used partially similar methods to those used in previous studies [5,12,16,19], but in contrast, we focused on the feasibility of objective assessments in daily life, with a task that could be conducted anywhere in less than 30 seconds using one's own smartphone. The smartphone is always with you, and a gamified task does not draw attention, even in public places. The low burden enables regular monitoring, providing continuous data of symptoms over time to support in treatment decisions.

Second, can the effects of PD medication be detected using the same inertial data captured during game sessions played before and after medication intake? In other words, we explored

whether or not medication-induced changes in motor symptoms could be measured using frequency-domain features extracted from accelerometer data. We classified the games played into two groups: "before" and "after" medication intake. For participants suffering from rigidity and bradykinesia, we found a statistically significant difference in particular signal characterizing features (Table 5). It is known that bradykinesia is usually responsive to PD medication [29]. Kostikis et al [14] also compared "off" and "on" medication states in laboratory conditions with two volunteers with PD. Even though they did not measure rigidity, according to the physician observing the measurements, the patients' rigidity improved after medication intake. Hence, our results are in line with their observations. Further research is needed to reproduce these findings and to investigate why we could not find a before-and-after difference for all participants and for tremor symptoms. To this end, we hypothesize that the time window of measurement could have had an effect on the results. It is possible that our participants were consistently under the effects of their medication, thus resulting in similar data across all game sessions.

Reflection on Smartphone-Based Monitoring of PD in Real Life

As this study was not a laboratory-controlled experiment, the way participants played the game could have affected the accelerometer data signal. For example, if the participant's hand was extended, such a position might have induced postural tremor, or if the arm was resting on their lap, rest tremor might have become dominant. Table 2 categorized the game sessions according to the F0 as dyskinesia, rest tremor, postural tremor, or kinetic tremor, as determined in a study by Pierleoni et al [5]. Indeed, based on the interview data reported in our previous study [21], some participants implemented strategies to "beat the game." For example, P10 and P12 reported to press their elbow against their torso to keep their hands steady, and P01 would occasionally play the game while holding the smartphone with both hands. These three participants had the most games classified as rest tremor (Table 2). P08 mentioned that the game was "easier" if sitting down, and a majority of his games were grouped as postural tremor. P07 noticed how his posture impacted the game and held his hand in such a way that it wasn't supported by his body; very likely as a consequence, 80% of his games were within the F0 postural tremor frequency area.

It should be noted that the F0 in tremor frequencies does not indicate tremor (Table 2). F0 indicates the frequency of maximum power in PSD but does not otherwise take into account the magnitude of the peak. The TIP describes the severity of the tremor effect, and with the F0 we can further characterize the type of tremor.

Limitations

With the STOP app, tremor analysis is limited to tremor severity in participants' hands, measured using their own smartphone as the instrument. Conversely, in assessments by health professionals, tools such as the UPDRS can be used to evaluate other tremor characteristics, such as amplitude in the legs, jaw, and neck. In our study, the fragmentation of the smartphone device base already caused minor issues, and this can only be expected to exacerbate in the future. To this end, measures to also track and account for the exact device make and model should be added to the approach.

Levodopa treatment is prescribed to alleviate motor symptoms, and although we know the medication intake time, we often ignore the symptoms in particular participants that the medication was prescribed for; hence, the analysis of the medication intake effect is preliminary (it is unclear whether or not the medication was meant to reduce tremor severity). Additionally, the time difference between game sessions and medication intake ("before" and "after" game sessions) varied, as did the magnitude of the changes induced by the medication, which were recorded by the STOP app. This time difference should be taken into account in future studies.

The participant sample size was admittedly small. However, this was compensated for by the high number of individual contributions in the form of game sessions. Further, our data analysis focused on results that generalize sufficiently well for the purposes of this paper: investigating the role of accelerometer data in differentiating between different symptoms and the effects of medication.

Future Work

Further research is needed to assess the internal and external validity of the TIP, as our results suggest it has the potential to quantify tremor severity. Previous studies [5,14,16,19] have based their tremor severity evaluations on the UPDRS, which evaluates the tremor on a scale from 0 to 4. Similarly, we used the self-assessed UPDRS II as a baseline in our measurements. However, the UPDRS was designed not for daily symptom severity assessment but rather to detect changes in symptom level in the long term. In future, we shall explore different alternatives to quantify tremor that we can use for baseline. One option is to utilize self-reporting about tremor severity, or to compare the TIP of a game session to the user's long-term average to provide insight into the variation in personal tremor severity levels.

Changes in our accelerometer features were inconsistent between "before" and "after" medication game sessions. This suggests that we could explore personalized tremor classification models. The effect of hand tremor is visible in the accelerometer signal, but we did not find a statistically significant effect of medication. In addition, it is necessary to focus on particular medication types and PD symptoms to explore the difference between "before" and "after" medication game sessions using accelerometer data in more homogeneous conditions.

Given the availability and sensing capabilities of smartphones, we envision that tools such as the STOP app can support the care and monitoring of PD as well as enable frequent, or even continuous, measuring of medication effects in naturalistic conditions. Even though real-life assessments pose a challenge for data quality due to differences in sensing devices and conditions, standalone smartphone solutions can have a lower burden, thus increasing engagement. For clinicians, a richer picture of symptom severity enabled by sensor data could enable them to better understand people's conditions and prescribe tailored medications.

Conclusions

In summary, it is feasible to detect and quantify the severity of hand tremor using accelerometer data collected with modern, off-the-shelf smartphones. We replicated and validated previously reported features derived from accelerometer data collected in real-world settings. To this end, we presented the TIP, a metric that could support further research into unobtrusive tremor assessment with smartphones but requires further internal and external validation. Additionally, we identified a statistically significant difference between the game sessions before and after medication intake among participants with rigidity and bradykinesia, and concluded that detecting the effects of PD medication is possible but further research is warranted.

Acknowledgments

This work is partially funded by the Academy of Finland (Grants 313224-STOP, 316253-SENSATE, 320089-SENSATE, 318927-6Genesis Flagship, and 318930-GenZ), and by personal research grants awarded by the Finnish Parkinson Foundation, the Tauno Tönning Foundation, the Jenny and Antti Wihuri Foundation, the Nokia Foundation, and the Emil Aaltonen Foundation.

The author Florian Wolling has been supported by the University of Siegen and the German Academic Exchange Service (DAAD), which enabled his research visit at the University of Oulu, Finland, Biomimetics and Intelligent Systems Group.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Participants' details.
[PDF File (Adobe PDF File), 152 KB - mhealth_v8i11e21543_app1.pdf]

Multimedia Appendix 2
Technical details of the linear interpolation of the accelerometer signal and the Welch method.
[PDF File (Adobe PDF File), 191 KB - mhealth_v8i11e21543_app2.pdf]

Multimedia Appendix 3
Wilcoxon rank sum test details of group comparisons.
[PDF File (Adobe PDF File), 155 KB - mhealth_v8i11e21543_app3.pdf]

References

1. Dorsey ER, Constantinescu R, Thompson JP, Biglan KM, Holloway RG, Kieburtz K, et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. Neurology 2006 Nov 02;68(5):384-386. [doi: 10.1212/01.wnl.0000247740.47667.03]
2. Carolinas P. What is Parkinson’s Disease?. 2020. URL: https://www.parkinsonassociation.org/what-is-parkinsons-disease/ [accessed 2020-11-06]
3. Politis M, Wu K, Molloy S, G. Bain P, Chaudhuri KR, Piccini P. Parkinson's disease symptoms: The patient's perspective. Mov. Disord 2010 May 14;25(11):1646-1651. [doi: 10.1002/mds.23135]
4. Bhatia KP, Bain P, Bajaj N, Elble RJ, Hallett M, Louis ED, Tremor Task Force of the International ParkinsonMovement Disorder Society. Consensus Statement on the classification of tremors. from the task force on tremor of the International Parkinson and Movement Disorder Society. Mov Disord 2018 Jan 30;33(1):75-87 [FREE Full text] [doi: 10.1002/mds.27121] [Medline: 29193359]
5. Pierleoni P, Palma L, Belli A, Pernini L. A real-time system to aid clinical classification and quantification of tremor in Parkinson's disease. 2014 Presented at: IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI). IEEE; 2014; Valencia p. 113-116. [doi: 10.1109/bhi.2014.6864317]
6. Kraus PH, Lemke MR, Reichmann H. Kinetic tremor in Parkinson’s disease – an underrated symptom. J Neural Transm 2006 Jan 30;113(7):845-853. [doi: 10.1007/s00702-005-0354-9]
7. Bötzel K, Tronnier V, Gasser T. The differential diagnosis and treatment of tremor. Dtsch Arztebl Int 2014 Mar 28;111(13):225-35; quiz 236 [FREE Full text] [doi: 10.3238/arztebl.2014.0225] [Medline: 24739887]
8. Raethjen J, Austermann K, Witt K, Zeuner KE, Papengut F, Deuschl G. Provocation of Parkinsonian tremor. Mov. Disord 2008 Mar 28;23(7):1019-1023. [doi: 10.1002/mds.22014]
9. Goetz C, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, Movement Disorder Society UPDRS Revision Task Force. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Movement disorders: official journal of the Movement Disorder Society 2008 Nov 15;23(15):2129-2170. [doi: 10.1002/mds.22340] [Medline: 19025984]

10. Zesiewicz TA, Sullivan KL, Hauser RA. Levodopa-induced dyskinesia in Parkinson's disease: epidemiology, etiology, and treatment. Curr Neurol Neurosci Rep 2007 Jul 12;7(4):302-310. [doi: 10.1007/s11910-007-0046-y] [Medline: 17618536]
11. LeMoyne R, Mastroianni T, Cozza M, Coroian C, Grundfest WS. Implementation of an iPhone for characterizing Parkinson's disease tremor through a wireless accelerometer application. 2010 Presented at: Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE. . IEEE; 31 Aug.-4 Sept. 2010; Buenos Aires p. 4954-4958. [doi: 10.1109/iembs.2010.5627240]
12. Barrantes S, Sánchez Egea AJ, González Rojas HA, Martí MJ, Compta Y, Valldeoriola F, et al. Differential diagnosis between Parkinson's disease and essential tremor using the smartphone's accelerometer. PLoS ONE 2017 Aug 25;12(8):e0183843. [doi: 10.1371/journal.pone.0183843]
13. Woods AM, Nowostawski M, Franz EA, Purvis M. Parkinson's disease and essential tremor classification on mobile device. Pervasive and Mobile Computing 2014 Aug;13:1-12. [doi: 10.1016/j.pmcj.2013.10.002]
14. Kostikis N, Hristu-Varsakelis D, Arnaoutoglou M, Kotsavasiloglou C. A Smartphone-Based Tool for Assessing Parkinsonian Hand Tremor. IEEE J. Biomed. Health Inform 2015 Nov;19(6):1835-1842. [doi: 10.1109/jbhi.2015.2471093]
15. Duque JDL, González-Vargas AM, Sánchez Egea AJ, González Rojas HA. Using Machine Learning and Accelerometry Data for Differential Diagnosis of Parkinson's Disease and Essential Tremor. 2019 Presented at: Workshop on Engineering Applications; 2019; Santa Marta, Colombia p. 368-378. [doi: 10.1007/978-3-030-31019-6_32]
16. Bazgir O, Habibi SH, Palma L, Pierleoni P, Nafees S. A classification system for assessment and home monitoring of tremor in patients with Parkinson's disease. J Med Signals Sens 2018;8(2):65. [doi: 10.4103/2228-7477.232088]
17. Bazgir O, Frounchi J, Habibi SAH, Palma L, Pierleoni P. A neural network system for diagnosis and assessment of tremor in Parkinson disease patients. 2015 Presented at: 22nd Iranian Conference on Biomedical Engineering (ICBME). . IEEE; 2015; Tehran, Iran p. 1-5. [doi: 10.1109/icbme.2015.7404105]
18. Locatelli P, Alimonti D. Differentiating essential tremor and Parkinson's disease using a wearable sensor — A pilot study. 2017 Presented at: 2017 7th IEEE International Workshop on Advances in Sensors and Interfaces; 2017; Vieste, Italy p. 213-218. [doi: 10.1109/iwasi.2017.7974254]
19. Pierleoni P, Belli A, Bazgir O, Maurizi L, Paniccia M, Palma L. A Smart Inertial System for 24h Monitoring and Classification of Tremor and Freezing of Gait in Parkinson's Disease. IEEE Sensors J 2019 Dec 1;19(23):11612-11623. [doi: 10.1109/jsen.2019.2932584]
20. Bermeo A, Bravo M, Huerta M, Soto A. A system to monitor tremors in patients with Parkinson's disease. 2016 Presented at: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2016; Orlando, FL, USA p. 5007-5010. [doi: 10.1109/embc.2016.7591852]
21. Kuosmanen E, Kan V, Vega J, Visuri A, Nishiyama Y, Dey AK, et al. Challenges of Parkinson's Disease: User Experiences with STOP. 2019 Presented at: the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services; October, 2019; Taipei, Taiwan p. 1-11. [doi: 10.1145/3338286.3340133]
22. Kuosmanen E, Kan V, Visuri A, Vega J, Nishiyama Y, Dey AK, et al. Mobile-based Monitoring of Parkinson's Disease. 2018 Presented at: the 17th International Conference on Mobile and Ubiquitous Multimedia; November, 2018; Cairo, Egypt p. 441-448. [doi: 10.1145/3282894.3289737]
23. Kan V. STOP: A smartphone-based game for Parkinson's disease medication adherence. Oulu: University of Oulu, Faculty of Information Technology and Electrical Engineering; 2018.
24. Android Developers. Motion sensors. 2020. URL: https://developer.android.com/guide/topics/sensors/sensors_motion [accessed 2020-11-06]
25. Loper E, Bird S. NLTK: the natural language toolkit. 2002 Presented at: ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics; 2002; Philadelphia. [doi: 10.3115/1118108.1118117]
26. Grosset KA, Bone I, Grosset DG. Suboptimal medication adherence in Parkinson's disease. Mov Disord 2005 Nov;20(11):1502-1507. [doi: 10.1002/mds.20602]
27. Welch P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Trans. Audio Electroacoust 1967 Jun;15(2):70-73. [doi: 10.1109/tau.1967.1161901]
28. Quora. Is it possible that the confidence interval is a negative number? Is it possible that the confidence interval is a negative number?. 2020. URL: https://www.quora.com/Is-it-possible-that-the-confidence-interval-is-a-negative-number [accessed 2020-06-11]
29. Burkhard PR, Shale H, Langston JW, Tetrud JW. Quantification of dyskinesia in Parkinson's disease: validation of a novel instrumental method. Mov Disord 1999 Sep;14(5):754-763. [doi: 10.1002/1531-8257(199909)14:5<754::aid-mds1007>3.0.co;2-1] [Medline: 10495036]

Abbreviations

AUC: area under the curve
F0: fundamental frequency
F50: central frequency

|F50-F0|: difference between F50 and F0
PD: Parkinson disease
PSD: power spectral density
PV: peak value
SF50: frequency dispersion
STOP: Sentient Tracking of Parkinson
TIP: tremor intensity parameter
UPDRS: Unified Parkinson Disease Rating Scale

Edited by G Eysenbach, L Buis; submitted 17.06.20; peer-reviewed by J Hernández-Vara, M Bardus; comments to author 27.07.20; revised version received 21.09.20; accepted 25.10.20; published 26.11.20.

Please cite as:
Kuosmanen E, Wolling F, Vega J, Kan V, Nishiyama Y, Harper S, Van Laerhoven K, Hosio S, Ferreira D
Smartphone-Based Monitoring of Parkinson Disease: Quasi-Experimental Study to Quantify Hand Tremor Severity and Medication Effectiveness
JMIR Mhealth Uhealth 2020;8(11):e21543
URL: http://mhealth.jmir.org/2020/11/e21543/
doi:10.2196/21543
PMID:33242017

Original Paper

A Mobile-Based Intervention for Dietary Behavior and Physical Activity Change in Individuals at High Risk for Type 2 Diabetes Mellitus: Randomized Controlled Trial

Zidu Xu[1,2], BSc; Ji Geng[2], BSc; Shuai Zhang[2], BSc; Kexin Zhang[3], BSc; Lin Yang[1], PhD; Jing Li[2*], PhD; Jiao Li[1*], PhD

[1]Institute of Medical Information and Library, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[2]School of Nursing, Peking Union Medical College, Beijing, China

[3]Nursing Department, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Jing Li, PhD
School of Nursing
Peking Union Medical College
No 33, Badachu Road, Shijingshan District
Beijing, 100144
China
Phone: 86 1088771124
Email: annelee13@126.com

Abstract

Background: Intensive lifestyle modifications have proved effective in preventing type 2 diabetes mellitus (T2DM), yet the efficiency and effectiveness of these modifications need to be improved. Emerging social media interventions are considered useful in promoting these lifestyles; nevertheless, few studies have investigated the effectiveness of combining them with behavior theory.

Objective: This study aims to examine the effectiveness of a 6-month mobile-based intervention (DHealthBar, a WeChat applet) combined with behavioral theory compared with a printed intervention in improving dietary behaviors, physical activity, and intention to change these behaviors among populations at high risk for T2DM.

Methods: Participants aged 23 to 67 years were recruited offline in Beijing, China, and were randomized into the intervention group or the control group, which received educational content via DHealthBar or a printed handbook, respectively. Educational materials were culturally tailored recommendations on improving dietary behaviors, physical activity, and intention to change based on the transtheoretical model. Participants in the intervention arm received push notifications twice per week on WeChat and had access to the educational content for the 6-month study period. Participants in the control arm received the same intervention content through printed materials. The outcomes of participants' behavior change, intention to change behavior, and anthropometric characteristics were collected via online measuring tools at baseline, 3 months, and 6 months.

Results: In this study, 79 enrolled individuals completed baseline information collection (control: n=38 vs intervention: n=41), and 96% (76/79) completed the 6-month follow-up visit. Attrition rates did not differ significantly between the 2 groups (χ^2_1=0.0, *P*=.61). Baseline equivalence was found. Participants in both groups reported a statistically significant decrease in energy intake at the 2 follow-up assessments compared with baseline (3 months, control: exp[β]=0.83, 95% CI 0.74-0.92 vs intervention: exp[β]=0.76, 95% CI 0.68-0.85; 6 months, control: exp[β]=0.87, 95% CI 0.78-0.96 vs intervention: exp[β]=0.57, 95% CI 0.51-0.64). At 6 months, a significantly larger decrease was observed in the intervention group in energy, fat, and carbohydrate intake, accompanied with a significantly larger increase in moderate-intensity physical activity compared with the control group (energy: exp[β]=0.66, 95% CI 0.56-0.77; fat: exp[β]=0.71, 95% CI 0.54-0.95; carbohydrates: exp[β]=0.83, 95% CI 0.66-1.03; moderate-intensity physical activity: exp[β]=2.05, 95% CI 1.23-3.44). After 6 months of the intervention, participants in the intervention group were more likely to be at higher stages of dietary behaviors (exp[β]=26.80, 95% CI 3.51-204.91) and physical activity (exp[β]=15.60, 95% CI 2.67-91.04) than the control group.

Conclusions: DHealthBar was initially effective in improving dietary behavior, physical activity, and intention to change these behaviors among populations who were at high risk of developing T2DM, with significant differences in the changes of outcomes over the 6-month intervention period.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2000032323; https://tinyurl.com/y4h8q4uf

(JMIR Mhealth Uhealth 2020;8(11):e19869) doi:10.2196/19869

KEYWORDS

transtheoretical model; type 2 diabetes mellitus; high risk; social media; dietary behavior; physical activity

Introduction

Type 2 diabetes mellitus (T2DM) accounts for over 90% of diabetes mellitus cases [1]. According to the International Diabetes Federation [2], China has the highest number of diabetic adults, with 1 in 10 adults aged 20 to 79 years (116 million adults) having diabetes mellitus. Unhealthy lifestyles are responsible for the increased risk of T2DM, including habitual sedentary behavior, overindulgence in animal fat, refined carbohydrates, and sugar-sweetened beverages [3-5]. Medication interventions, such as metformin, have been recommended as a preventive measure; nevertheless, for long-term effects, lifestyle interventions provide better performance, as described in follow-up studies such as the US Diabetes Prevention Program (DPP) and the Da Qing Study [6,7]. The 30-year Da Qing follow-up study reported that a median delay in diabetes onset of 3.96 years and a reduction in diabetes incidence and its complications were achieved, with an additional 1.44-year extension in life expectancy in the intervention group [6]. However, despite these promising results in curbing the diabetes epidemic, the typical face-to-face approaches of behavioral interventions might have trouble reaching a wider population to achieve and sustain healthy lifestyles, especially considering the high cost of current procedures (ie, physician visits, nutrition services, personal consultants, etc) [7].

With advanced information and communication technology and ubiquitous smartphones, medical services based on mobile phones and social networking services have been suggested as promising alternatives for cost-effective delivery of scalable behavior change interventions [8]. Trials have shown the potential of social media (eg, WeChat, Facebook, and Twitter) in delivering behavioral interventions on weight loss, diet pattern reconstruction, and online exercise coaching [9-11]. WeChat, a Chinese multipurpose social media app with a wide range of functions, had nearly 1.151 billion monthly active users on the service platform and over 300 million daily active users of extending applets in the third quarter of 2019, making it well suited to disseminate mobile health (mHealth) interventions in China [12,13]. However, the explosion of mHealth interventions and services does not mean that these interventions are ready for wide use. Rather, they have yet to be sufficiently explored and evaluated. If mHealth interventions were purely self-guided, a lot of information would overload the individuals and deter their adherence [14]. Participants might end up not using the intervention, resulting in low response.

Investigators have called for incorporating behavior theory to nudge mHealth interventions to be more effective in promoting health behaviors [15]. The transtheoretical model (TTM) has long been the most widely accepted model of health behavior change [16]. TTM focuses on the intention of individuals and assumes that the decision making and implementation of behavior change is a continuous, cyclical process divided into precontemplation, contemplation, preparation, action, and maintenance [17]. It has been suggested that TTM performs well in measuring one's current stage of behavior change and accounting for the dynamic process over time [18,19], thus making it possible to provide tailored interventions for individuals at various stages of health behavior change. The individuals would therefore be propelled toward higher stages of behavior change.

TTM-based behavior interventions have been proven effective [18-21]. Lee et al [22] noted that TTM-based interventions have been applied to physical activity and diet improvement among adolescents and male workers using web-based delivery methods. TTM-based interventions have also been successful in lifestyle management among patients with diabetes [22,23]. However, few studies have combined social media delivery modes with TTM-based strategies in tailored health behavior interventions, let alone applied them to high-risk populations for T2DM. In this context, we focused on high-risk individuals and used TTM-based interventions delivered through social media to expand the delivery modes and elements of the TTM-based interventions.

The aim of this study was to investigate the effectiveness of a 6-month WeChat-based (mobile-based) behavior intervention driven by TTM compared with a print-based intervention in promoting dietary behaviors and physical activity among individuals at high risk for T2DM. The secondary purpose was to compare changes in participants' intentions to change these behaviors. Changes in anthropometric characteristics (ie, BMI and waist circumference) were also considered.

Methods

Research Design

The effectiveness of the TTM-based social media intervention was examined with a 2-arm randomized controlled trial registered on the Chinese Clinical Trial Registry (ChiCTR2000032323). Under the condition of receiving a written and oral explanation of the program requirements and giving informed consent, participants were randomly allocated to either (1) the DHealthBar intervention group or (2) the control group, which received printed materials. The control group was set to obtain stand-alone information about the effect of theory-driven interventions in improving health behaviors

[18-20]. All study protocols, the consent process, and participant communications were approved by the School of Nursing, Peking Union Medical College institutional review board for the protection of human subjects (202002). See Multimedia Appendix 1 for the CONSORT-EHEALTH checklist for this trial.

Participant Recruitment and Eligibility Criteria

Individuals were eligible if they met the following inclusion criteria: (1) aged 18 years or older, (2) high risk for diabetes, as measured by the American Diabetes Association (ADA) screening tool (score of 5 or more) [24], (3) access to WeChat push notifications with a smartphone, and (4) agreement to informed consent and further participation in the study. Exclusion criteria for participants were (1) diagnosis of diabetes or other endocrine diseases (eg, thyroid disease under treatment), (2) mobility impairment, (3) diet therapy due to physical illness, (4) uncontrolled hypertension (systolic blood pressure ≥160 mmHg, diastolic blood pressure ≥110 mmHg), (5) participation in a lifestyle intervention program or similar study or a plan to take medicine or undergo surgery to lose weight, or (6) pregnancy. Elimination criteria for participants were withdrawal from the study voluntarily or failure to complete the follow-up due to an illness.

Participants (n=81) who met the inclusion criteria were recruited from May to June 2018 in the physical examination center of Peking Union Medical College Hospital and were randomly allocated to the intervention group (social media intervention) or control group (printed material intervention) using a randomization list generated by a computerized program. The enrolled participants were linked to the online baseline questionnaires on site after providing signed informed consent and voluntarily completing the baseline survey. They were given a detailed explanation by researchers to help them understand the questions and the study better, as well as a brief (10-minute) explanation that they were at high risk for T2DM and that improved dietary behaviors and physical activity could help them delay or prevent the incidence. This in-person interaction allowed related web links for further data collection and interventions to be received by participants immediately; the subsequent procedure could be accessible to enrolled participants later. Due to the characteristics and design of this study, participants were not blinded to group assignment; data analysts, however, were blinded to the assignment.

Study Setting and Data Collection

The enrolled participants were informed of group assignment by phone, text messages, or emails after leaving the study site. Both groups were provided with the same culturally tailored educational materials to encourage lifestyle modification, modified from the US DPP design plan and other evidence-based strategies [25]. TTM-based behavioral intervention techniques and strategies were applied in the delivery of the interventions to both groups. Participants in the mobile-based intervention group were told to follow the WeChat subscription account to access the details of the intervention program, including push notifications for health education content and links to online self-report questionnaires. Participants in the control group were mailed a lifestyle modification handbook. There was no direct interaction between participants in the different groups after randomization, and participants had no further contact with research staff except for reminders for the 3-month and 6-month online follow-up assessments. Reminder notifications for assessments were standardized across groups in content, delivery, and frequency to avoid bias from interactional differences. All follow-up assessment data were collected via WeChat applets connected to an encrypted database. Logic checks were conducted to identify suspicious submissions, and the suspected data errors were dropped from the collections and re-collected after participants understood the meaning of the questions with assistance from the research team members. To enhance the adherence and engagement of participants, each participant was compensated with office supplies valued at around ¥70 (US $10.50) per assessment.

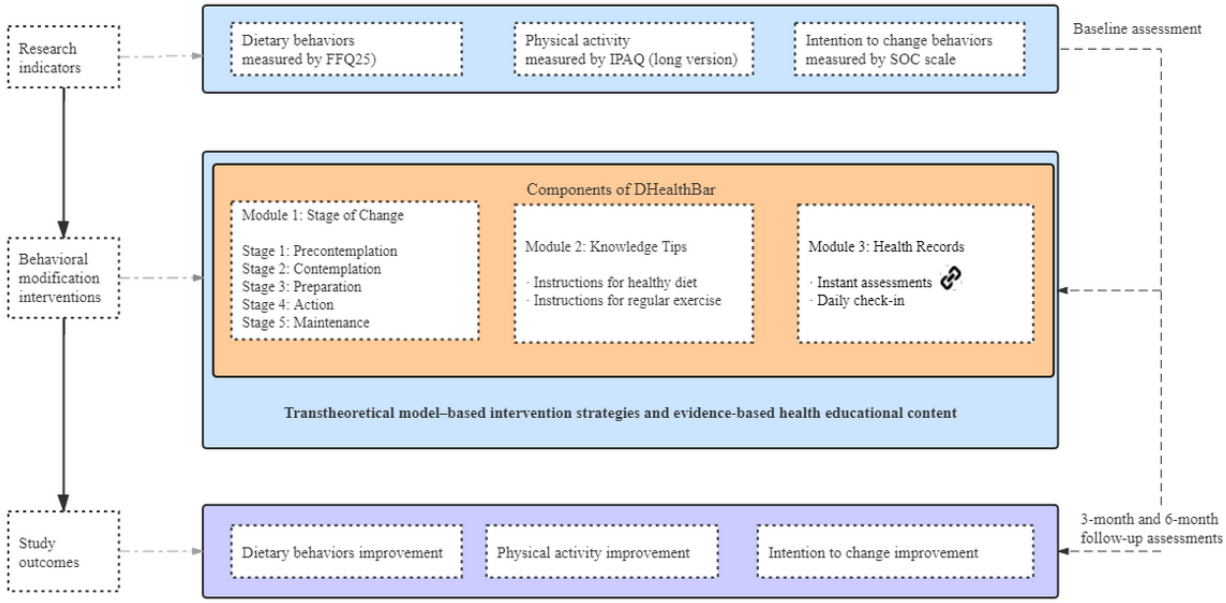
Study Conditions

The DHealthBar Intervention

DHealthBar provided 6 months of a WeChat-based behavior intervention aimed at improving eating habits and physical activity, thus reducing diabetes risk in the long term. It also aimed to increase the intention to change these behaviors. The 6-month mobile-based intervention was composed of educational material sent by the WeChat subscription account named DHealthBar, WeChat applets (lightweight apps that form part of the WeChat ecosystem, which could be used independently) embedded with online questionnaires, and a check-in applet serving as an online forum with functions similar to Twitter Moments.

The general design of DHealthBar and supporting details are displayed in Figure 1. It was designed to educate people at high risk for T2DM about diabetes prevention and specifically focus on providing practical strategies on relevant aspects, such as (1) interventions on behavior change, (2) behavior change instructions, (3) behavior change tracking tools (ie, online questionnaires), and (4) a common space for communication and sharing. Within the WeChat account, 3 modules were set to categorize the push messages: stages of change, knowledge tips, and health records (Figure 1). In the stages of change section, researchers created message interventions for the research participants to improve their motivation for behavior change (ie, from precontemplation to maintenance). The knowledge tips section provided practical instructions for participants to change their dietary behaviors and physical activity. In the health records section, the food frequency questionnaire 25 (FFQ25) diet-tracking tool [26], the International Physical Activity Questionnaire (IPAQ) physical activity tracking tool (long version in Chinese) [27], and the Stage of Change (SOC) Scale stages of change tracking tool [28] were presented in the instant assessment embedded in the WeChat applets and allowed immediate feedback upon submission. Furthermore, there was a check-in online forum for participants to record and share daily diet and exercises in the form of photos and texts; participants could appreciate and comment on tweets that interested them to encourage peers indirectly.

Figure 1. The schema of the TTM-based social media behavior intervention plan. FFQ25: food frequency questionnaire 25; IPAQ: International Physical Activity Questionnaire; SOC: Stage of Change; TTM: transtheoretical model.



With the support of the DPP design plan [25] and localized evidence-based guidelines [29,30], the push messages were designed and enhanced with TTM-based [31] and evidence-based social media lifestyle intervention strategies [10,32,33] to achieve behavior change goals.

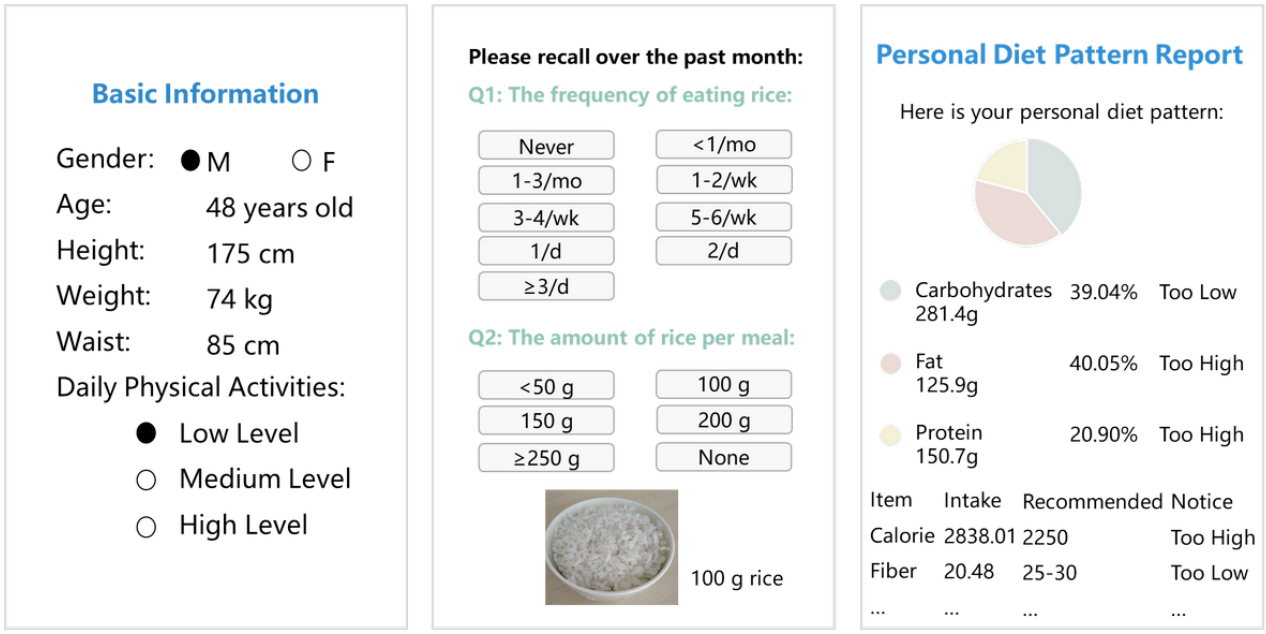
The educational content was designed following several principles. For eating behaviors, the focus was on decreasing added sugars and refined carbohydrates, decreasing saturated and trans fats, and increasing fruits and vegetables. Changes in food types and reduction in portion size were emphasized as means of reducing energy intake rather than specific calorie targets or calorie counting. For dietary behaviors, changing food types and reducing portion size, for instance, limiting the intake of sweetened beverages and refined carbohydrates, increasing fruits and vegetables, and replacing saturated and trans fats with polyunsaturated fats, was meant to (1) reduce total energy intake (in kcal per day) according to age, gender, and level of daily activities [30] and (2) modify the proportion of 3 macronutrients, with carbohydrates at 50% to 65%, protein at 10% to 15%, and fats at 20% to 30% [30].

For physical activity, participants were encouraged to enhance the endurance and frequency of their physical activity to achieve the goal of at least 30 minutes of moderate-intensity physical activity per day [25,29]. Resistance training was also encouraged.

The participant's psychological readiness to take action for these behavioral changes was addressed using TTM-based behavior change techniques, specifically following strategies aimed at behavior change at different stages [17,28]. For instructions on behavior change skills, the educational material was arranged into comics, stories, or short articles and then sent as push messages on the WeChat subscription. The participants could leave messages about topics of interest and share the recommended content as their WeChat Moments. All presented information was designed with input from an endocrinologist, a diabetes nurse educator, registered dietitians, and health behavior specialists and was reviewed by them before release. Educational content about dietary behaviors and physical activity were posted for all engaging participants twice a week (Multimedia Appendix 2).

WeChat mini applets embedded with online questionnaires were used as automated trackers to encourage participants to monitor their daily health behavior and increase engagement (Figure 2). After completing the assessments, participants received automated feedback and personalized recommendations immediately. In particular, the processes of behavior change were used as indicators for tailored interventions for the participant's intention to change and as a general guide for the behavior promotion intervention modifications in the next study period [31]. Due to the dynamics of intentions and lifestyle changes, an interval of 3 months was set to reevaluate the participants and rearrange the intervention domains and skills accordingly.

Figure 2. Screenshots of the online assessment for dietary behaviors. From left to right: the entry screen for basic information, the food frequency questionnaire 25, and the feedback information on personal dietary behaviors.



Alternative Care: Self-Learning With DPP-Based Handbook

The alternative care for participants in the control group was a behavior change handbook with instructions on dietary behaviors and physical activity modification. A follow-up parallel to the intervention group was conducted after 3 months and 6 months with online questionnaires embedded in the WeChat applets. Participants in the control group were informed of the assessment link at baseline assessment and received reminders through phone calls, text messages, or emails around the follow-up assessment time points. Apart from these time points, participants in the control group had no contact with researchers during the study period.

Outcome Measures

The primary outcome measures were changes in dietary behaviors and physical activity at the 3-month and 6-month follow-up from baseline, serving as the key goals of our interventions. Secondary outcome measures were (1) stage of change for dietary behaviors and physical activity, which reflects one's intention for behavior change, and (2) anthropometric characteristics, including BMI and waist circumference. Dietary behaviors were described using daily intake of energy; the macronutrients protein, carbohydrates, and fat; and macronutrient proportions. Physical activity was described using weekly physical activity in total and weekly physical activity of light, moderate, and vigorous intensity. Participants completed self-reported assessments at 3 time points: baseline, 3 months, and 6 months.

General Information

General information was collected during the screening process, including gender, age, body mass, body mass index, waist circumference, family history of diabetes, occupation, and education.

Diabetes Risk Screening

Diabetes self-screening tools published by the ADA for risk of type 2 diabetes [24] were adopted and included 6 measures—gender, age, family history, hypertension history, BMI, and sedentariness—with a total of 10 points. Scoring less than 5 points indicated a low risk of diabetes and scoring 5 or more points indicated the state of being prediabetic or at a high risk for diabetes.

Changes in Dietary Behaviors and Physical Activity

Changes in dietary behaviors and physical activity were evaluated to examine the effect of the interventions. The simplified FFQ25 developed by Gao et al [26] was adopted for the dietary behavior assessment. There are 25 types of staple foods and nonstaple foods included. The frequency and weight of food intake were recorded for calculation and analysis. The frequency of food intake was on a 9-point scale, ranging from "never eat" to "more than 3 times a day," with a corresponding weight sequence of 0, 0.03, 0.07, 0.22, 0.50, 0.79, 1, 2, and 3. The amount of food intake was on a 6-point scale, ranging from "never eat" to "more than 3 times a day," with a corresponding weight sequence of 0, 0.5, 0.75, 1, 1.5, and 2. The FFQ25 has been demonstrated to be a diet assessment tool with high reliability and validity. Results for dietary behaviors included energy intake (in kcal) per day, 3 macronutrient (ie, carbohydrates, fat, and protein) intakes (in g) per day, and macronutrient proportions (percentage), with higher values indicating greater intake of certain types of nutrients or corresponding types of food. Physical activity was assessed using the self-administered IPAQ (long version in Chinese) [27]. The IPAQ contains a total of 27 questions covering a comprehensive set of 4 domains of physical activity (ie, leisure time physical activity, domestic and gardening (yard) activities, work-related physical activity, and transport-related physical activity), with 2 added questions on sedentary behavior. In addition, items in the form were structured to provide details about the specific types of walking (*W*), moderate-intensity

activity (*M*), and vigorous-intensity activity (*V*) within each domain. The volume of activity measured by the IPAQ can be weighted by metabolic equivalents of task (METs) that represent the energy required to yield a score in MET minutes. With selected MET values of each type, the MET score was calculated as the specific MET value (*W*, *M*, *V*) × weekly frequency (days per week) × time per day (minutes per day). In this study, because of the equivalent intensity, we used light-intensity activity (*L*) to replace walking activity (*W*) as one of our outcomes. With the weekly light-, moderate-, and vigorous-intensity physical activity summed up, we could determine the MET score for the weekly total physical activity.

In addition to the standardized approach, we adopted long-form IPAQ data processing rules to perform data cleaning and set an upper and lower threshold for the duration of activity, thus qualifying the results within a reasonable range [27].

Stage of Behavior Change

The SOC Scale, derived from the SOC Scale for reducing dietary fat intake used by Greene and Rossi [28], was adopted to evaluate the intention to change dietary behaviors and physical activity. The Cronbach α coefficient for each item was no less than .76. The stage assessment was divided into 5 phases based on the answers from the participants (precontemplation: "Do not intend to make dietary/physical activity changes in the next 6 months"; contemplation: "Intend to make dietary/physical activity changes in the next 6 months"; preparation: "Intend to make dietary/physical activity changes in the next 30 days"; action: "Have been changing dietary/physical activity but it has been less than 6 months"; maintenance: "Have been changing dietary/physical activity for more than 6 months").

Statistics Analysis

Based on established methods [34,35], it was estimated that a sample of 74 participants would be required to detect a 715 kcal change in daily energy intake from baseline to 6 months with a standard deviation of 300 kcal, using an α level of .05 and a power level of 80%. The final desired estimated number of participants enrollment was set at 78 after 5% estimated attrition. Descriptive statistics were used to examine the baseline information, using number and percentage for nominal categorical variables and mean and standard deviation or median and interquartile range for normally or not normally distributed variables, respectively. Comparisons between groups at all 3 assessment points (baseline, 3 months, 6 months) were made on demographic measures (baseline only) and all outcome measures (ie, dietary behaviors, physical activity, stage of behavior change, and anthropometric characteristics). We used 2-tailed *t* tests for continuous variables that were normally distributed. Kruskal-Wallis tests were used for ordinal variables or continuous variables not normally distributed. Chi-square tests were used for nominal categorical variables.

To analyze the repeated outcome measures [36,37], generalized linear mixed models (GLMMs) were used to examine changes over time and differences between the 2 arms in dietary behaviors, physical activity, stage of behavior change, and anthropometric characteristics, which allowed response variables from different distributions, such as nonnormal distribution responses and ordinal categorical responses. Given both fixed and random effects, the general form of the model is [38]:

$$y=\beta X+uZ+\varepsilon$$

In the above equation, *y* is the outcome variable, *X* is a predictor variable, β is a fixed-effects regression coefficient, and *uZ* is designed for random effects, with *Z* being the random complement to the fixed *X* and *u* being the random complement to the fixed β. In addition, ε is designed for the residuals, the part of *y* that is not explained by the model.

The following form of the model is given to make this more concrete in this study setting:

$$y=\beta_1 X_1+\beta_2 X_2+\beta_3 X_1 X_2+Z+\varepsilon$$

In the above equation, *y* is the outcome variable, either continuous (ie, dietary behavior outcome measures, physical activity outcome measures, BMI, and waist circumference) or categorical (ie, stage of change for dietary behaviors or physical activity). $\beta_1X_1 + \beta_2X_2 + \beta_3X_1X_2$ represents fixed effects. Further, in this study, variables were included for intervention (X_1: 0=no, 1=yes) and time (X_2: 1=baseline, 2=3-month follow-up, 3=6-month follow-up), and βs are fixed-effects regression coefficients. Particularly, to interpret the unique effect of the intervention on time, an intervention-time interaction item ($\beta_3X_1X_2$) was added. *Z*, the random intercept, is incorporated for the random effects of participants. All analyses were adjusted for baseline age, gender, education level, and occupational level for the probability of an impact on outcome measures [39,40]. Outcomes of GLMM analyses were reported as exponentiated coefficients (exp[β]). Adjusted *P* values and confidence intervals to control the type I error rate at 0.05 (95% CI) were computed through simulations (Multimedia Appendices 3 and 4). Specific model type, link function, and the total number of observations included in the model are listed in the footnotes of Multimedia Appendices 3 and 4.

We conducted post hoc comparisons of the intervention group versus the control group at the 3-month follow-up, the intervention group versus the control group at the 6-month follow-up, the difference in mean changes or proportion changes between the intervention group and the control group at the 3-month and 6-month follow-up, and within-group comparisons between each time point (ie, 0 months vs 3 months, 0 months vs 6 months, 3 months vs 6 months). The first 2 post hoc comparisons evaluated whether the mean levels of outcome measures in the intervention group differed from those of the control group at each time point. The middle 2 comparisons evaluated whether there was a change in the differences between the treatment group and the control group over time. The last comparisons examined the changes in outcome measures across the study period. Besides these comparisons, for the main-effects-only model, if the intervention was statistically significant, post hoc comparisons were conducted to determine if outcome measures differed significantly in mean or proportion between the 2 groups.

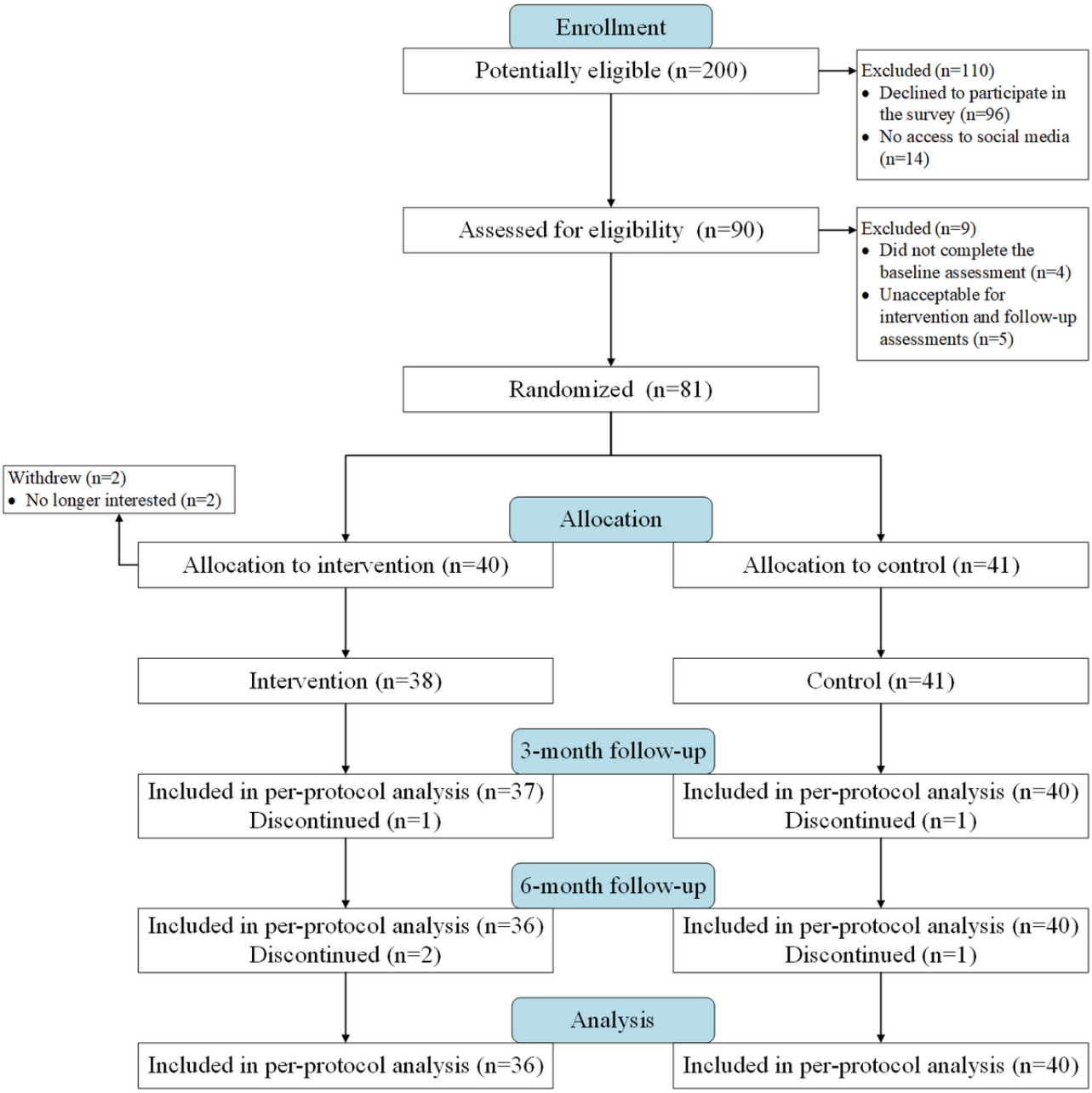
All the tests were 2-tailed and used a .05 significance level. All analyses were conducted using per-protocol analysis with R (version 3.6.2; R Foundation for Statistical Computing).

Results

Overview

A pragmatic, parallel-group, randomized controlled trial in Beijing, China, was conducted from May 2018 to March 2019. Of the 200 potentially eligible participants, 90 (45.0%) agreed to a further test. After excluding 7 participants who declined to participate further (2 after the informed group assignment) and 4 who provided invalid questionnaires, a total of 79 of the 90 (88%) participants were randomly allocated to the social media intervention group (n=38) or control group (n=41). There was high retention and participation in follow-up assessments; of the 79 participants, 78 (99%) were followed up with at 3 months and 76 (96%) were followed up with at 6 months. Analyses of the attrition rate did not show statistically significant difference between the 2 groups (χ^2_1=0.0; *P*=.61). Hence, attrition was not systemic. Figure 3 shows the participants' flow through the study.

Figure 3. Flow diagram for the randomized trial.



The participants were middle-aged (control group: median 47.5, IQR 44.0-55.0 years; intervention group: 46.0, IQR 44.0-48.2 years) and well educated, with 51 of the 76 participants (67%) having a college degree or higher. At baseline, most participants were at relatively low stages (ie, precontemplation, contemplation, or preparation) of dietary behavior or physical activity change. Baseline equivalence of participants was demonstrated between the 2 groups in sociodemographic, anthropometric, and behavioral characteristics (Table 1). No participants in the study did not complete the follow-up assessment for the reason of being diagnosed with diabetes, nor did anyone report adverse events caused by the social media intervention.

Table 1. Characteristics and baseline behaviors of participants in the DHealthBar trial.

Characteristics	Control (n=40)	Intervention (n=36)	*t* test[a] (*df*)	H value[b] (*df*)	Chi-square test (*df*)	*P* value
Age (years), median (IQR)	47.5 (44.0-55.0)	46.0 (44.0-48.2)	N/A[c]	2.03 (1)	N/A	.15
Gender, n (%)			N/A	N/A	0.4[d] (1)	.52
Male	24 (60.0)	18 (50.0)				
Female	16 (40.0)	18 (50.0)				
Body mass (kg), median (IQR)	70.0 (65.0-80.5)	72.7 (65.0-80.0)	N/A	0.55 (1)	N/A	.46
BMI (kg/m²), median (IQR)	24.7 (23.4-26.1)	25.3 (24.7-26.2)	N/A	2.75 (1)	N/A	.10
Waist circumference (cm), median (IQR)	81.0 (73.7-85.0)	82.5 (77.7-88.0)	N/A	1.48 (1)	N/A	.22
Family history of diabetes, n (%)			N/A	N/A	1.7[d] (1)	.19
Yes	13 (32.5)	18 (50.0)				
No	27 (67.5)	18 (50.0)				
Occupational classification, n (%)			N/A	N/A	1.9[e] (9)	.99
Administrator or manager	3 (7.5)	2 (5.5)				
Office and administrative support	12 (30.0)	10 (28.0)				
Health care practitioner	2 (5.0)	2 (5.5)				
Legal profession	3 (7.5)	2 (5.5)				
Architect or engineer	3 (7.5)	5 (14.0)				
Laborer or protective service worker	4 (10.0)	4 (11.1)				
Driver	3 (7.5)	2 (5.5)				
Teacher	4 (10.0)	4 (11.1)				
Computer programmer	2 (5.0)	3 (8.3)				
Retired	4 (10.0)	2 (5.5)				
Education level, n (%)			N/A	N/A	0.9[e] (2)	.72
Primary	1 (2.5)	1 (2.8)				
Secondary	14 (35.0)	9 (25.0)				
Tertiary	25 (62.5)	26 (72.2)				
Dietary behaviors (FFQ25[f])						
Energy intake (kcal/d), median (IQR)	2029.0 (1515.0-2547.0)	2263.0 (2047.0-2419.0)	N/A	2.12 (1)	N/A	.14
Macronutrients intake (g/d), median (IQR)						
Fat	78.3 (56.9-103.2)	81.2 (30.9-38.3)	N/A	0.001 (1)	N/A	.97
Carbohydrates	210.7 (141.3-272.4)	228.8 (153.2-256.3)	N/A	0.29 (1)	N/A	.59
Protein	94.7 (80.4-128.5)	98.3 (63.2-123.9)	N/A	0.78 (1)	N/A	.38
Macronutrients proportion (%)						
Fat, median (IQR)	37.3 (33.0-40.9)	36.8 (30.9-38.3)	N/A	1.18 (1)	N/A	.28
Carbohydrates, mean (SD)	42.7 (7.2)	44.6 (6.2)	1.22 (73.88)	N/A	N/A	.23
Protein, mean (SD)	20.8 (2.3)	20.1 (2.4)	–1.20 (72.03)	N/A	N/A	.23
Physical activity (IPAQ [g]), median MET[h]/wk (IQR)						
Total	2632.5 (1577.7-3746.2)	2385.5 (803.2-3852.0)	N/A	0.81 (1)	N/A	.37

Characteristics	Control (n=40)	Intervention (n=36)	*t* test[a] (*df*)	H value[b] (*df*)	Chi-square test (*df*)	*P* value
Light intensity	891.0 (334.1-1980.0)	627.0 (457.9-1485.0)	N/A	0.49 (1)	N/A	.48
Moderate intensity	1230.5 (612.5-2520.5)	720.0 (315.0-1590.0)	N/A	2.08 (1)	N/A	.15
Vigorous intensity	0.0 (0.0-495.0)	0.0 (0.0-740.0)	N/A	0.8 (1)	N/A	.37
Stage of dietary behaviors change (SOC[i]), n (%)			N/A	0.14 (1)	N/A	.71
Precontemplation	13 (32.5)	7 (19.4)				
Contemplation	15 (37.5)	20 (55.6)				
Preparation	6 (15.0)	6 (16.7)				
Action	3 (7.5)	1 (2.8)				
Maintenance	3 (7.5)	2 (5.5)				
Stage of physical activity change (SOC), n (%)			N/A	0.02 (1)	N/A	.89
Precontemplation	7 (17.5)	6 (16.7)				
Contemplation	22 (55.0)	19 (52.8)				
Preparation	5 (12.5)	6 (16.7)				
Action	4 (10.0)	5 (13.8)				
Maintenance	2 (5.0)	0 (0.0)				

[a]2-sample, 2-tailed *t* test.

[b]Kruskal-Wallis test.

[c]N/A: not applicable.

[d]Chi-squared test.

[e]Fisher exact test.

[f]FFQ25: simplified food frequency questionnaire 25.

[g]IPAQ: International Physical Activity Questionnaire (long version in Chinese).

[h]MET: metabolic equivalent of task.

[i]SOC: Stage of Change.

Primary Outcomes: Dietary Behaviors and Physical Activity Changes

The results of the outcome measures and univariate analyses at the 3- and 6-month follow-up assessments are reported in Multimedia Appendix 5. For outcome measures in dietary behaviors, mean levels of energy intake (*P*<.001), macronutrient intake of fat and protein (fat: *P*<.001; protein: *P*<.001), and proportion of fat (*P*<.001) were significantly lower in the intervention group than in the control group at the 6-month follow-up assessment, while the proportion of carbohydrates was significantly higher than that in the control group (*P*<.001) (Multimedia Appendix 5).

Significant group × time interaction effects and significant main effects for time were both observed in intake of energy and the 3 macronutrients, as well as in total physical activity and moderate-intensity physical activity per week (Table 2 and Figure S1 in Multimedia Appendix 6). Significant main effects for time were also observed in light-intensity physical activity per week. Significant main effects for the intervention were only observed in the intake of protein and carbohydrate proportions (Table 2).

For dietary behaviors, energy intake decreased significantly at 3 months and 6 months compared with baseline in both groups (control: 3 months, *P*<.001; 6 months, *P*=.003; intervention: 3 months, *P*<.001; 6 months, *P*<.001). See Multimedia Appendix 3 for exp(β) and 95% CI values. The intake of the 3 macronutrients decreased significantly at 6 months compared with baseline in both groups (control: fat, *P*=.047; carbohydrates, *P*=.01; protein, *P*=.006; intervention: fat, *P*<.001; carbohydrates, *P*<.001; protein, *P*=.001). Intake of fat (*P*=.049) and protein (*P*<.001) decreased significantly at 3 months compared with baseline in the control group. A significant decrease in fat proportion was observed in the intervention group at 6 months (*P*=.01). For physical activity, total physical activity and light-intensity physical activity per week improved significantly at 6 months compared with baseline in both groups (control: total, *P*=.03; light-intensity physical activity, *P*=.01; intervention: total, *P*<.001; light-intensity physical activity, *P*<.001). There were continuously significant changes in energy intake (decrease) and moderate-intensity physical activity (increase) in the intervention group (3 vs 0 months: energy, *P*<.001; moderate-intensity physical activity, *P*=.04; 6 vs 3 months: energy, *P*<.001; moderate-intensity physical activity, *P*=.03). *P* values in this paragraph were adjusted.

At 6 months, the intake of energy (*P*<.001), fat (*P*<.001), and protein (*P*=.003) and the intake proportion of fat (*P*=.003) were significantly lower in the intervention group than in the control group, while the proportion of carbohydrate intake (*P*=.03) was significantly higher in the intervention group than in the control group. See Multimedia Appendix 4 for exp(β) and 95% CI

values. Significantly larger decreases over 6 months were observed in the intervention group in energy (P<.001), fat (P<.001), and carbohydrate (P<.001) intakes compared with the control group. In addition, the increase in moderate-intensity physical activity was significantly larger in the intervention group than in the control group (P=.002). P values in this paragraph were adjusted.

Table 2. The effect of the intervention and time on primary and secondary outcomes over the intervention period.

Outcome measures	Model effects					
	Group		Time		Group × time	
	F test (*df*)	*P* value	*F* test (*df*)	*P* value	*F* test (*df*)	*P* value
Dietary behaviors (FFQ25[a])						
Energy intake[b]	1.07 (1,70)	.30	73.30 (2,148)	<.001[c]	27.59 (2,148)	<.001[c]
Macronutrient intake[b]						
Fat	3.23 (1,70)	.08	23.26 (2,148)	<.001[c]	7.07 (2,148)	.001[c]
Carbohydrates	0.08 (1,70)	.78	21.78 (2,148)	<.001[c]	4.58 (2,148)	.01[c]
Protein	4.83 (1,70)	.03[c]	24.73 (2,148)	<.001[c]	5.88 (2,148)	.003[c]
Macronutrients proportion (%)						
Fat[b]	3.91 (1,70)	.05	2.01 (2,148)	.14	2.56 (2,148)	.08
Carbohydrates[b]	4.75 (1,70)	.03[c]	1.90 (2,148)	.15	1.65 (2,148)	.19
Protein[d]	2.67 (1,70)	.11	1.76 (2,148)	.17	0.16 (2,148)	.85
Physical activity (IPAQ[e])						
Total[d]	0.04 (1,70)	.84	19.38 (2,148)	<.001[c]	3.68 (2,148)	.03[c]
Light intensity[f]	0.39 (1,66.83)	.53	19.42 (2,133)	<.001[j]	1.77 (2,133)	.17
Moderate intensity[g]	0.16 (1,66.65)	.69	9.63 (2,139.7)	.001[c]	6.84 (2,139.7)	.001[c]
Vigorous intensity[h]	3.87 (1,50.53)	.05	0.08 (2,76.37)	.92	0.51 (2,76.07)	.60
Stage of dietary behavior change (SOC[i,j])	3.56 (1,145)	.06	6.37 (2,145)	.002[c]	27.18 (2,145)	<.001[c]
Stage of physical activity change (SOC[i])	2.66 (1,145)	.10	6.31 (2,145)	.002[c]	47.76 (2,145)	<.001[c]
Anthropometric characteristics[b]						
BMI (kg/m²)	0.47 (1,70)	.50	125.40 (2,148)	<.001[c]	21.68 (2,148)	<.001[c]
Waist circumference (cm)	4.54 (1,70)	.04[c]	172.47 (2,148)	<.001[c]	44.62 (2,148)	<.001[c]

[a]FFQ25: simplified food frequency questionnaire 25.

[b]Model (response variable of gamma distribution with log link) included age, gender, education level, and occupational classification as covariates. Number of observations=228.

[c]P values represent statistically significant results, P<.05.

[d]Model (response variable of lognormal distribution with identity link) included age, gender, education level, and occupational classification as covariates. Number of observations=228.

[e]IPAQ: International Physical Activity Questionnaire (long version in Chinese).

[f]Model (response variable of lognormal distribution with identity link) included age, gender, education level, and occupational classification as covariates. Number of observations=213.

[g]Model (response variable of lognormal distribution with identity link) included age, gender, education level, and occupational classification as covariates. Number of observations=220.

[h]Model (response variable of lognormal distribution with identity link) included age, gender, education level, and occupational classification as covariates. Number of observations=132.

[i]Model (response variable of multinomial distribution with cumulative logit link) included age, gender, education level, and occupational classification as covariates. Number of observations=228.

[j]SOC: Stage of Change.

Secondary Outcomes: Stage of Behavior Change and Anthropometric Characteristics

According to univariate analyses (Multimedia Appendix 5), there was a statistically significant difference between the 2 groups in the stage of change for dietary behaviors and physical activity at 6 months *(P*<.001) (Multimedia Appendix 5). A statistically significant difference between the 2 groups in the stage of change for dietary behaviors was also observed at 3 months (*P*=.02) (Multimedia Appendix 5).

Significant group × time interaction effects and significant main effects for time were both observed at 6 months in the stage of change for dietary behaviors and physical activity, as well as in anthropometric characteristics (ie, BMI and waist circumference) (Table 2 and Figure S2 in Multimedia Appendix 6). Significant main effects of the intervention were only observed in anthropometric characteristics (Table 2).

There were continuous, statistically significant changes in all secondary outcome measures in the intervention group (3 vs 0 months: diet SOC, *P*<.001; physical activity SOC, *P*<.001; BMI, *P*<.001; waist, *P*<.001; 6 vs 3 months: diet SOC, *P*<.001, physical activity SOC, *P*<.001; BMI, *P*<.001; waist, *P*<.001). See Multimedia Appendix 3 for exp(β) and 95% CI values. Statistically significant changes in all secondary outcome measures were observed in the control group at 6 months (diet SOC: *P*<.001; physical activity SOC: *P*<.001; BMI: *P*<.001; waist: *P*<.001). *P* values in this paragraph were adjusted.

There were significant between-group differences in the distribution of the stages of change for dietary behaviors and physical activity at 6 months; participants in the intervention group had a higher probability of being at higher stages for these behavior changes (diet SOC: *P*=.002; physical activity SOC: *P*=.01). See Multimedia Appendix 4 for exp(β) and 95% CI values. Significantly greater improvements were observed at 6 months in the probability of being at a higher stage of change in dietary behaviors (*P*=.002) and physical activity (*P*=.003). There was also a significantly greater decrease in anthropometric characteristics in the intervention group (BMI: *P*<.001; waist: *P*<.001). Additionally, the decrease in waist circumference was significantly greater in the intervention group than in the control group at 3 months (*P*<.001). Adjusted *P* values were used in this paragraph.

Discussion

Principal Findings

This study examined the effectiveness of DHealthBar in improving diet behaviors and physical activity for people at high risk for T2DM. TTM-based behavioral change techniques were applied during the research. Our findings show that this WeChat applet could be effective in inducing positive dietary changes and improving physical activity within a 6-month intervention period. Significant decreases in energy and macronutrient intake were found in both groups, but these decreases differed substantially between groups in energy, fat, and carbohydrate intake. In addition, considerable improvements over 6 months in total and light-intensity physical activity were found in both groups, with no substantial between-group differences. This social media intervention also showed positive effects on participants' intentions to change these behaviors (measured using stages of behavior change) and anthropometric characteristics.

These improved outcomes are comparable to those of other mobile-based intervention studies that have shown favorable results in improving dietary behaviors and physical activity [33,40-42]. Regardless of the limitations of self-reported assessments, it was encouraging to find a continuous change in dietary behaviors and physical activity associated with the social media intervention. Furthermore, a high level of user engagement and retention was achieved within 6 months. We owe these benefits partly to the merits of the WeChat platform, which is wide reaching, cost-efficient, and easy to use. Shared by disparate populations, WeChat enables health educators to present complex content in the form of vivid graphics or videos, thus spreading health education content for the public in a feasible and accessible way [43]. Given the heavy burden resulting from the growing prevalence of T2DM and the fact that high-risk populations with sedentary lifestyles and dietary patterns rich in fat and refined carbohydrates are much more likely to develop diabetes and other chronic diseases associated with unhealthy lifestyles [1,4], it is plausible to adopt mobile-based approaches to facilitate the provision of behavioral interventions and increase the uptake with good scalability [5,44]. As such, our findings might be insightful for population-level delivery of diet change and physical activity improvement techniques for diabetes prevention, given the wide reach of social media interventions [5,43].

The significant decreases in daily energy and macronutrient intakes in both groups indicated that interventions incorporating behavior change theory and techniques tend to be effective. This observation was supported by previous randomized studies based on TTM behavior intervention strategies [19,20,41]. Changes in dietary macronutrient patterns were examined in this study, and a significant decrease in the proportion of fat intake was observed in the intervention group. Previous studies targeting dietary change have focused on the reduction of energy intake or the consumption of certain foods, such as vegetables, fruits, and low-fat milk [40,41,45]. Interestingly, to the best of our knowledge, few studies have focused on the effect of mHealth interventions on changes in macronutrient patterns. Therefore, this finding might be useful, as it suggests the potential of mHealth interventions to modify dietary macronutrient patterns, which have proved critical to weight loss and cardiovascular risk management [46].

Designed as an extension of DPP-based lifestyle interventions, DHealthBar provided the participants in the intervention group with behavioral support after baseline that was delivered through social media, whereas those in the control group underwent a completely self-guided behavior modification. There would be an enhancement in the efficacy of mHealth or eHealth interventions if they were provided with behavioral support or guidance from counselors either online or face to face, with an insight shared in conclusions related to diabetes prevention and other areas [47]. Consistent with this, a continuous improvement in dietary behaviors and physical activity was observed. In the control group, however, considerable heterogeneity existed in

the effects over time, and the participants tended to relapse or stagnate after an immediate improvement (<3 months). DHealthBar, which was delivered through a social media platform, provided interaction between care providers and receivers. Since the push notifications with lifestyle instructions were available to all the participants in the intervention group, participants could comment below them. Researchers could be informed of the participant's preferences and requests for educational content and improve the content accordingly as soon as possible.

Aligned with traditional mHealth interventions delivered through text messages, websites, or smartphone apps [33,40,42,48], behavior change techniques, including self-monitoring, goal setting, and tailored educational content and feedback, may have contributed to the effectiveness of the dietary change and physical activity intervention in this study. As a result of the self-report tools embedded in the WeChat platform, individuals could receive timely automated feedback to better understand their stage achievements and current status [48,49], thereby enhancing their engagement. In addition, some of the effects may have been due to the strengths inherent in social media, especially the sense of interpersonal connection, which is underlined as a key factor of the success of mHealth programs [50]. In the DHealthBar trial, a sense of community was created. Participants in the intervention arm were connected via a social network, where clustered peers concerned about the same health issue provided social and emotional support to each other in an online forum [11,41]. Additionally, they were encouraged to track their daily routines of food or exercise and post them as "moments" to receive appreciation and comments from others.

The positive effects of the social media intervention on physical activity were modest. A significant difference between the 2 groups, however, was observed in moderate-intensity physical activity improvement after 6 months, which is partly ascribed to sociodemographic factors. Most participants in this study were middle-aged and had relatively high socioeconomic indicators. The physical activity of middle-aged or older adults shares a strong pattern of high frequency and slow pace [51]. Accordingly, the exercise program recommended in this study tended to include items such as brisk walking, tai chi, and Chinese square dancing, whereas vigorous physical activity and muscle strength training were only mentioned occasionally. Despite the evidence that adults at higher socioeconomic levels are more likely to show greater engagement in physical activity, due to their improved health-related knowledge and skills [39,52], the type of activity was limited.

One of the strengths of this study was the theory-driven nature of the study design, which allowed the procedures of the intervention and evaluation to be relatively empirical. TTM-based behavioral intervention techniques sensitively measured the stages of change that participants underwent and precisely identified the fluctuations in participants' intentions to change behaviors that were affected by dynamical barriers or facilitators. Hierarchical strategies could therefore be used according to the readiness of individuals [53]. Additionally, the progression of participants' behavior improvement in this study kept pace with participants' stage of change level. One explanation is that substantial change in actual behavior might partly be ascribed to interventions concerned with cognition and intention modification, which serve as critical mediators of actual behavioral change [54]. As previously claimed, the degree of motivation among participants is likely to influence results [52,55].

In this study, changes in anthropometric markers (ie, BMI and waist circumference decreases) were additional benefits in the context of an intensive behavior modification trial. This finding corresponds with the results from similar studies [41,56,57]. To maintain relatively low monetary and time costs, we included physiological measures, such as BMI and waist circumference, but not biochemical markers of physiological health. However, given the genetic and somatotype characteristics of diabetes in East Asia [58,59], as well as the probable independence of glycemic benefit from weight loss or behavior modification [40,41], it is plausible to further include biochemical markers associated with glycemic improvements to induce more indicative clinical outcomes [60].

We acknowledge several limitations of this study. First, our sample is limited in terms of both size and geographical coverage, which might attenuate the power and generalizability of our study. Second, due to initial financial resource limitations and the desire to maintain relatively low attrition by assuring participants' active engagement within a reasonable period, this randomized trial lasted for only 6 months. Third, participants were not blinded to randomization; consequently, we could not exclude the social desirability bias, that is, whether the postassessments or stand-alone interventions acted as incentives to the participants to provide a positive response. Fourth, due to considerations of cost-effectiveness and convenience, self-reported questionnaires were used for data collection in this study [39], which achieved the additional goal of avoiding the potential burden resulting from wearable tracking tools [45]. Given the characteristics of our assessment tools, which had high reliability and validity, and the characteristics of our participants, who had high education levels, the confounding effect might be insignificant [61]. However, there is no denying that recall bias did exist, which may have resulted in an erroneous estimation of the association between behavior changes and interventions, as mentioned by previous studies [61,62]. Last, despite the strong evidence base for the TTM, substantial heterogeneity remains in the effects of health behavior interventions based on this model, which have not been confirmed uniformly [31,63].

In this study, we demonstrated the initial effect of DHealthBar, a mobile behavior intervention coupled with theory-driven instructions, on individuals recruited from a single region. Investigators in future studies should assess the effectiveness in more diverse socioeconomic contexts. In addition, the usage of this social media platform could be measured to obtain more insights into the participants' perceptions of the intervention program. Furthermore, novel measurement tools could be included in the future to objectively measure the outcomes, under the condition of such tools being affordable and acceptable. Analyses of images and text on social media with health behavior tags might help health professionals obtain a better understanding of the modification procedure that the participants underwent and reveal potential clinical influences

[64,65]. Future work should elucidate the “active ingredients” of the techniques, as well as effective ways to employ them [66], thus shedding light on the design of mHealth interventions.

Conclusion

DHealthBar was effective in promoting dietary behaviors and physical activity among populations at high risk for T2DM. Our findings, coupled with high engagement and retention, may suggest the strength of social media as a health behavior intervention delivery mode, inspiring further work on diet and physical activity promotion.

Acknowledgments

This research was supported by the Peking Union Medical College National College Students' innovation and entrepreneurship training program (2018zlgc0626), the China Medical Board Nursing Faculty Research Grant (13-168-201611), and the Chinese Academy of Medical Sciences (grant No. 2018-I2M-AI-016 and 2018PT33024).

The study required the efforts of a multidisciplinary research and development team. The authors would like to thank Professor Yanping Liu for providing nutrition consultation and Dr Liang Xu for providing computer expertise. The authors would like to thank all the participants in this study. The authors would also like to thank Peking Union Medical College Hospital for approving the participant recruitment and the provision of the venue.

Authors' Contributions

ZX, Jing L, JG, SZ, and KZ conceived of and designed the study. ZX performed the statistical analysis and authored the manuscript with input from JG, SZ, KZ, Jing L, LY, and Jiao L. Jiao L conducted the workflow design and data analysis, LY and ZX worked on the data integration, and KZ assisted with data preprocessing and analysis. ZX, Jing L, Jiao L, JG, SZ, and KZ conducted the research procedure. Jing L contributed considerably to the theoretical framework construction and coordinated and directed all aspects of study conduction. ZX promoted the development of programs, supported the cloud server construction, and contributed to web design. Jiao L contributed to the DHealthBar workflow design and mHealth-related data analysis. JG, SZ, and KZ contributed significantly to providing behavioral improvement support for participants and offered suggestions for program development and data processing. ZX, JG, SZ, and KZ maintained the social media platform and collected data. Jing L, LY, and Jiao L supervised ZX, reviewed and revised the manuscript with valuable suggestions, and gave ZX oversight of the coaching and training. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2131 KB - mhealth_v8i11e19869_app1.pdf]

Multimedia Appendix 2
TTM-based behavioral intervention strategies and contents.
[PDF File (Adobe PDF File), 95 KB - mhealth_v8i11e19869_app2.pdf]

Multimedia Appendix 3
Within-group comparisons of outcomes over the intervention period.
[PDF File (Adobe PDF File), 83 KB - mhealth_v8i11e19869_app3.pdf]

Multimedia Appendix 4
Between-group comparisons of outcomes over the intervention period.
[PDF File (Adobe PDF File), 165 KB - mhealth_v8i11e19869_app4.pdf]

Multimedia Appendix 5
Results for outcome measures at 3-month and 6-months follow-up assessments and univariate analyses.
[PDF File (Adobe PDF File), 119 KB - mhealth_v8i11e19869_app5.pdf]

Multimedia Appendix 6
Interaction plots of outcome measures.
[PDF File (Adobe PDF File), 214 KB - mhealth_v8i11e19869_app6.pdf]

References

1. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nat Rev Endocrinol 2018 Feb;14(2):88-98. [doi: 10.1038/nrendo.2017.151] [Medline: 29219149]
2. International Diabetes Federation. China Country report 2010-2045. IDF Diabetes Atlas, 9th ed. 2019. URL: https://www.diabetesatlas.org/data/en/country/42/cn.html [accessed 2020-02-18]
3. Xu JY, Dan QH, Chan V, Wat NMS, Tam S, Tiu SC, et al. Genetic and clinical characteristics of maturity-onset diabetes of the young in Chinese patients. Eur J Hum Genet 2005 Apr;13(4):422-427 [FREE Full text] [doi: 10.1038/sj.ejhg.5201347] [Medline: 15657605]
4. Ma RCW, Lin X, Jia W. Causes of type 2 diabetes in China. Lancet Diabetes Endocrinol 2014 Dec;2(12):980-991. [doi: 10.1016/S2213-8587(14)70145-7] [Medline: 25218727]
5. Chan JCN, Zhang Y, Ning G. Diabetes in China: a societal solution for a personal challenge. Lancet Diabetes Endocrinol 2014 Dec;2(12):969-979. [doi: 10.1016/S2213-8587(14)70144-5] [Medline: 25218728]
6. Gong Q, Zhang P, Wang J, Ma J, An Y, Chen Y, Da Qing Diabetes Prevention Study Group. Morbidity and mortality after lifestyle intervention for people with impaired glucose tolerance: 30-year results of the Da Qing Diabetes Prevention Outcome Study. Lancet Diabetes Endocrinol 2019 Jun;7(6):452-461. [doi: 10.1016/S2213-8587(19)30093-2] [Medline: 31036503]
7. Ely EK, Gruss SM, Luman ET, Gregg EW, Ali MK, Nhim K, et al. A National Effort to Prevent Type 2 Diabetes: Participant-Level Evaluation of CDC's National Diabetes Prevention Program. Diabetes Care 2017 Oct;40(10):1331-1341 [FREE Full text] [doi: 10.2337/dc16-2099] [Medline: 28500215]
8. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009 Apr;30(1):273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
9. Hales S, Turner-McGrievy GM, Wilcox S, Fahim A, Davis RE, Huhns M, et al. Social networks for improving healthy weight loss behaviors for overweight and obese adults: A randomized clinical trial of the social pounds off digitally (Social POD) mobile app. Int J Med Inform 2016 Oct;94:81-90. [doi: 10.1016/j.ijmedinf.2016.07.003] [Medline: 27573315]
10. Kankanhalli A, Shin J, Oh H. Mobile-Based Interventions for Dietary Behavior Change and Health Outcomes: Scoping Review. JMIR Mhealth Uhealth 2019 Jan 21;7(1):e11312 [FREE Full text] [doi: 10.2196/11312] [Medline: 30664461]
11. He C, Wu S, Zhao Y, Li Z, Zhang Y, Le J, et al. Social Media-Promoted Weight Loss Among an Occupational Population: Cohort Study Using a WeChat Mobile Phone App-Based Campaign. J Med Internet Res 2017 Oct 23;19(10):e357 [FREE Full text] [doi: 10.2196/jmir.7861] [Medline: 29061555]
12. CIW Team. WeChat statistical highlights 2020; mini program DAU >300m. China Internet Watch. 2020 May 18. URL: https://www.chinainternetwatch.com/30201/wechat-stats-2019/ [accessed 2020-07-18]
13. Zhang X, Wen D, Liang J, Lei J. How the public uses social media wechat to obtain health information in china: a survey study. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):66 [FREE Full text] [doi: 10.1186/s12911-017-0470-0] [Medline: 28699549]
14. Crisp D, Griffiths K. Participating in online mental health interventions: who is most likely to sign up and why? Depress Res Treat 2014;2014:790457 [FREE Full text] [doi: 10.1155/2014/790457] [Medline: 24804089]
15. 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]
16. Glanz K, Rimer BK, Marcus LF. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass; 1990.
17. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
18. Menezes MCD, Mingoti SA, Cardoso CS, Mendonça RD, Lopes ACS. Intervention based on Transtheoretical Model promotes anthropometric and nutritional improvements - a randomized controlled trial. Eat Behav 2015 Apr;17:37-44. [doi: 10.1016/j.eatbeh.2014.12.007] [Medline: 25553558]
19. de Freitas PP, de Menezes MC, Dos Santos LC, Pimenta AM, Ferreira AVM, Lopes ACS. The transtheoretical model is an effective weight management intervention: a randomized controlled trial. BMC Public Health 2020 May 11;20(1):652 [FREE Full text] [doi: 10.1186/s12889-020-08796-1] [Medline: 32393214]
20. Spencer L, Adams TB, Malone S, Roy L, Yost E. Applying the transtheoretical model to exercise: a systematic and comprehensive review of the literature. Health Promot Pract 2006 Oct;7(4):428-443. [doi: 10.1177/1524839905278900] [Medline: 16840769]
21. Hall KL, Rossi JS. Meta-analytic examination of the strong and weak principles across 48 health behaviors. Prev Med 2008 Mar;46(3):266-274. [doi: 10.1016/j.ypmed.2007.11.006] [Medline: 18242667]
22. Lee JY, Park H, Min YH. Transtheoretical Model-based Nursing Intervention on Lifestyle Change: A Review Focused on Intervention Delivery Methods. Asian Nurs Res (Korean Soc Nurs Sci) 2015 Jun;9(2):158-167 [FREE Full text] [doi: 10.1016/j.anr.2015.05.001] [Medline: 26160246]

23. Jackson R, Asimakopoulou K, Scammell A. Assessment of the transtheoretical model as used by dietitians in promoting physical activity in people with type 2 diabetes. J Hum Nutr Diet 2007 Feb;20(1):27-36. [doi: 10.1111/j.1365-277X.2007.00746.x] [Medline: 17241190]
24. Risk Test. American Diabetes Association. URL: https://www.diabetes.org/risk-test [accessed 2018-03-05]
25. The Diabetes Prevention Program Research Group. The Diabetes Prevention Program. Design and methods for a clinical trial in the prevention of type 2 diabetes. Diabetes Care 1999 Apr;22(4):623-634. [doi: 10.2337/diacare.22.4.623] [Medline: 10189543]
26. Gao J, Fei J, Jiang L. Assessment of the reproducibility and validity of a simple food-frequency questionnaire used in diet. Acta Nutrimenta Sinica 2011;33(5):452-456. [doi: 10.13325/j.cnki.acta.nutr.sin.2011.05.012]
27. Fan M, Lyu J, He P. Chinese guidelines for data processing and analysis concerning the International Physical Activity Questionnaire. Chin J Epidemiol 2014 Aug;35(8):961-964 [FREE Full text]
28. Greene GW, Rossi SR. Stages of change for reducing dietary fat intake over 18 months. J Am Diet Assoc 1998 May;98(5):529-534. [doi: 10.1016/S0002-8223(98)00120-5] [Medline: 9597025]
29. Chinese Diabetes Society. Guidelines for the prevention and control of type 2 diabetes in China (2017 Edition). Chin J Diabetes Mellitus 2018 Jan;10(1):4-67. [doi: 10.3760/cma.j.issn.1674-5809.2018.01.003]
30. The Chinese Nutrition Society. Dietary Guidelines for Chinese Residents. Beijing, China: People's Medical Publishing House; 2016.
31. Prochaska JO, Redding CA, Evers KE. The transtheoretical model and stages of change. In: Prochaska JO, Redding CA, Evers KE, editors. Health behavior: Theory, research, and practice. 5th ed. New York, NY: Wiley Press; 2015:125-148.
32. Fraccaro P, Lavery-Blackie S, Van der Veer SN, Peek N. Behavioural Phenotyping of Daily Activities Relevant to Social Functioning Based on Smartphone-Collected Geolocation Data. Stud Health Technol Inform 2019 Aug 21;264:945-949. [doi: 10.3233/SHTI190363] [Medline: 31438063]
33. Pfammatter A, Spring B, Saligram N, Davé R, Gowda A, Blais L, et al. mHealth Intervention to Improve Diabetes Risk Behaviors in India: A Prospective, Parallel Group Cohort Study. J Med Internet Res 2016 Dec 05;18(8):e207 [FREE Full text] [doi: 10.2196/jmir.5712] [Medline: 27496271]
34. Wittes J. Sample size calculations for randomized controlled trials. Epidemiol Rev 2002;24(1):39-53. [doi: 10.1093/epirev/24.1.39] [Medline: 12119854]
35. Cohen JE. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 1988.
36. Detry MA, Ma Y. Analyzing Repeated Measurements Using Mixed Models. JAMA 2016 Jan 26;315(4):407-408. [doi: 10.1001/jama.2015.19394] [Medline: 26813213]
37. Liski EP. Generalized Linear Mixed Models: Modern Concepts, Methods and Applications by Walter W. Stroup. Int Stat Rev 2013 Nov;81(3):482-483. [doi: 10.1111/insr.12042_24]
38. Stroup WW, Milliken GA, Claassen EA, Wolfinger RD. SAS for Mixed Models: Introduction and Basic Applications. Cary, NC: SAS Institute; 2018.
39. Brindal E, Freyne J, Saunders I, Berkovsky S, Smith G, Noakes M. Features predicting weight loss in overweight or obese participants in a web-based intervention: randomized trial. J Med Internet Res 2012 Dec 12;14(6):e173 [FREE Full text] [doi: 10.2196/jmir.2156] [Medline: 23234759]
40. Duncan M, Vandelanotte C, Kolt GS, Rosenkranz RR, Caperchione CM, George ES, et al. Effectiveness of a web- and mobile phone-based intervention to promote physical activity and healthy eating in middle-aged males: randomized controlled trial of the ManUp study. J Med Internet Res 2014 Jun;16(6):e136 [FREE Full text] [doi: 10.2196/jmir.3107] [Medline: 24927299]
41. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes Prevention and Weight Loss with a Fully Automated Behavioral Intervention by Email, Web, and Mobile Phone: A Randomized Controlled Trial Among Persons with Prediabetes. J Med Internet Res 2015 Oct 23;17(10):e240 [FREE Full text] [doi: 10.2196/jmir.4897] [Medline: 26499966]
42. Ek A, Alexandrou C, Söderström E, Bergman P, Delisle Nyström C, Direito A, et al. Effectiveness of a 3-Month Mobile Phone-Based Behavior Change Program on Active Transportation and Physical Activity in Adults: Randomized Controlled Trial. JMIR Mhealth Uhealth 2020 Jun 08;8(6):e18531 [FREE Full text] [doi: 10.2196/18531] [Medline: 32510462]
43. Zhang X, Wen D, Liang J, Lei J. How the public uses social media wechat to obtain health information in china: a survey study. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):66 [FREE Full text] [doi: 10.1186/s12911-017-0470-0] [Medline: 28699549]
44. Cefalu WT, Petersen MP, Ratner RE. Response to comment on Cefalu et Al. The alarming and rising costs of diabetes and prediabetes: a call for action! Diabetes care 2014;37:3137-3138. Diabetes Care 2015 May;38(5):e82-e83 [FREE Full text] [doi: 10.2337/dc15-0215] [Medline: 25908165]
45. Mummah S, Robinson TN, Mathur M, Farzinkhou S, Sutton S, Gardner CD. Effect of a mobile app intervention on vegetable consumption in overweight adults: a randomized controlled trial. Int J Behav Nutr Phys Act 2017 Sep 15;14(1):125 [FREE Full text] [doi: 10.1186/s12966-017-0563-2] [Medline: 28915825]

46. Ge L, Sadeghirad B, Ball GDC, da Costa BR, Hitchcock CL, Svendrovski A, et al. Comparison of dietary macronutrient patterns of 14 popular named dietary programmes for weight and cardiovascular risk factor reduction in adults: systematic review and network meta-analysis of randomised trials. BMJ 2020 Apr 01;369:m696. [doi: 10.1136/bmj.m696] [Medline: 32238384]
47. Joiner KL, Nam S, Whittemore R. Lifestyle interventions based on the diabetes prevention program delivered via eHealth: A systematic review and meta-analysis. Prev Med 2017 Jul;100:194-207 [FREE Full text] [doi: 10.1016/j.ypmed.2017.04.033] [Medline: 28456513]
48. Lisón JF, Palomar G, Mensorio MS, Baños RM, Cebolla-Martí A, Botella C, et al. Impact of a Web-Based Exercise and Nutritional Education Intervention in Patients Who Are Obese With Hypertension: Randomized Wait-List Controlled Trial. J Med Internet Res 2020 Apr 14;22(4):e14196 [FREE Full text] [doi: 10.2196/14196] [Medline: 32286232]
49. Griauzde D, Kullgren JT, Liestenfeltz B, Ansari T, Johnson EH, Fedewa A, et al. A Mobile Phone-Based Program to Promote Healthy Behaviors Among Adults With Prediabetes Who Declined Participation in Free Diabetes Prevention Programs: Mixed-Methods Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Jan 09;7(1):e11267 [FREE Full text] [doi: 10.2196/11267] [Medline: 30626566]
50. Alley SJ, Schoeppe S, Rebar AL, Hayman M, Vandelanotte C. Age differences in physical activity intentions and implementation intention preferences. J Behav Med 2018 Jun;41(3):406-415. [doi: 10.1007/s10865-017-9899-y] [Medline: 29116569]
51. Lin Y, Yan J. Muscle-Strengthening Activities and Sociodemographic Correlates among Adults: Findings from Samples in Mainland China. Int J Environ Res Public Health 2020 Mar 27;17(7):2266 [FREE Full text] [doi: 10.3390/ijerph17072266] [Medline: 32230937]
52. Sheeran P, Maki A, Montanaro E, Avishai-Yitshak A, Bryan A, Klein WMP, et al. The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: A meta-analysis. Health Psychol 2016 Nov;35(11):1178-1188. [doi: 10.1037/hea0000387] [Medline: 27280365]
53. Khaylis A, Yiaslas T, Bergstrom J, Gore-Felton C. A review of efficacious technology-based weight-loss interventions: five key components. Telemed J E Health 2010 Nov;16(9):931-938 [FREE Full text] [doi: 10.1089/tmj.2010.0065] [Medline: 21091286]
54. Katan MB. Weight-loss diets for the prevention and treatment of obesity. N Engl J Med 2009 Feb 26;360(9):923-925. [doi: 10.1056/NEJMe0810291] [Medline: 19246365]
55. Balk-Møller C, Poulsen SK, Larsen TM. Effect of a Nine-Month Web- and App-Based Workplace Intervention to Promote Healthy Lifestyle and Weight Loss for Employees in the Social Welfare and Health Care Sector: A Randomized Controlled Trial. J Med Internet Res 2017 Apr 10;19(4):e108 [FREE Full text] [doi: 10.2196/jmir.6196] [Medline: 28396303]
56. Seo DC, Niu J. Evaluation of Internet-Based Interventions on Waist Circumference Reduction: A Meta-Analysis. J Med Internet Res 2015 Jul 21;17(7):e181 [FREE Full text] [doi: 10.2196/jmir.3921] [Medline: 26199208]
57. Kodama S, Saito K, Tanaka S, Horikawa C, Fujiwara K, Hirasawa R, et al. Effect of Web-based lifestyle modification on weight control: a meta-analysis. Int J Obes (Lond) 2012 May;36(5):675-685. [doi: 10.1038/ijo.2011.121] [Medline: 21694698]
58. Zimmet P, Shi Z, El-Osta A, Ji L. Epidemic T2DM, early development and epigenetics: implications of the Chinese Famine. Nat Rev Endocrinol 2018 Dec;14(12):738-746. [doi: 10.1038/s41574-018-0106-1] [Medline: 30310152]
59. Kim DS, Kim BC, Daily JW, Park S. High genetic risk scores for impaired insulin secretory capacity doubles the risk for type 2 diabetes in Asians and is exacerbated by Western-type diets. Diabetes Metab Res Rev 2018 Jan;34(1):e2944. [doi: 10.1002/dmrr.2944] [Medline: 29048714]
60. Zhao L, Zhang F, Ding X, Wu G, Lam YY, Wang X, et al. Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. Science 2018 Mar 09;359(6380):1151-1156. [doi: 10.1126/science.aao5774] [Medline: 29590046]
61. Krall EA, Dwyer JT, Ann Coleman K. Factors influencing accuracy of dietary recall. Nutrition Research 1988 Jul 19;8(7):829-841. [doi: 10.1016/S0271-5317(88)80162-3]
62. Prince SA, Cardilli L, Reed JL, Saunders TJ, Kite C, Douillette K, et al. A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2020 Mar 04;17(1):31 [FREE Full text] [doi: 10.1186/s12966-020-00938-3] [Medline: 32131845]
63. Mastellos N, Gunn LH, Felix LM, Car J, Majeed A. Transtheoretical model stages of change for dietary and physical exercise modification in weight loss management for overweight and obese adults. Cochrane Database Syst Rev 2014 Feb 05(2):CD008066. [doi: 10.1002/14651858.CD008066.pub3] [Medline: 24500864]
64. Wilkinson JL, Strickling K, Payne HE, Jensen KC, West JH. Evaluation of Diet-Related Infographics on Pinterest for Use of Behavior Change Theories: A Content Analysis. JMIR Mhealth Uhealth 2016 Dec 08;4(4):e133 [FREE Full text] [doi: 10.2196/mhealth.6367] [Medline: 27932316]
65. Pila E, Mond JM, Griffiths S, Mitchison D, Murray SB. A thematic content analysis of #cheatmeal images on social media: Characterizing an emerging dietary trend. Int J Eat Disord 2017 Jun;50(6):698-706. [doi: 10.1002/eat.22671] [Medline: 28075492]

66. Simeon R, Dewidar O, Trawin J, Duench S, Manson H, Pardo Pardo J, et al. Behavior Change Techniques Included in Reports of Social Media Interventions for Promoting Health Behaviors in Adults: Content Analysis Within a Systematic Review. J Med Internet Res 2020 Jun 11;22(6):e16002 [FREE Full text] [doi: 10.2196/16002] [Medline: 32525482]

Abbreviations

ADA: American Diabetes Association
DPP: Diabetes Prevention Program
FFQ25: food frequency questionnaire 25
GLMM: generalized linear mixed model
IPAQ: International Physical Activity Questionnaire
MET: metabolic equivalent of task
mHealth: mobile health
SOC: Stage of Change
T2DM: type 2 diabetes mellitus
TTM: transtheoretical model

Edited by G Eysenbach; submitted 05.05.20; peer-reviewed by H Wu, Z He, A Videira-Silva; comments to author 27.05.20; revised version received 22.07.20; accepted 02.10.20; published 03.11.20.

Please cite as:
Xu Z, Geng J, Zhang S, Zhang K, Yang L, Li J, Li J
A Mobile-Based Intervention for Dietary Behavior and Physical Activity Change in Individuals at High Risk for Type 2 Diabetes Mellitus: Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(11):e19869
URL: https://mhealth.jmir.org/2020/11/e19869
doi:10.2196/19869
PMID:33141092

Original Paper

Patient Perception of Mobile Phone Apps for the Care and Prevention of Sexually Transmitted Diseases: Cross-Sectional Study

Lena Jakob[1], MD; Theresa Steeb[1,2,3], MPH; Zeno Fiocco[1], MD; Teodora Pumnea[1], MD; Sophia Nomi Jakob[4], MSc; Anja Wessely[1,2,3], MSc; Christoph Clemens Rothenberger[1], MD; Titus Josef Brinker[5,6], MD; Lars Einar French[2], MD; Carola Berking[1,2,3], MD; Markus Vincent Heppt[1,2,3], MD, MSc, MHBA

[1]Department of Dermatology and Allergy, University Hospital Munich, Ludwig Maximilian University Munich, Munich, Germany

[2]Department of Dermatology, Universitätsklinikum Erlangen, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany

[3]Deutsches Zentrum für Immuntherapie, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany

[4]Arbeitsstelle für Diagnostik und Evaluation, Johann Wolfgang Goethe University Frankfurt, Frankfurt, Germany

[5]National Center for Tumor Diseases, German Cancer Research Center, Heidelberg, Germany

[6]Department of Dermatology, University Hospital Heidelberg, Heidelberg, Germany

Corresponding Author:
Markus Vincent Heppt, MD, MSc, MHBA
Department of Dermatology
Universitätsklinikum Erlangen
Friedrich-Alexander-University Erlangen-Nürnberg
Ulmenweg 18
Erlangen, 91054
Germany
Phone: 49 913185 ext 45810
Email: markus.heppt@uk-erlangen.de

Abstract

Background: In the emerging era of digitalization and electronic health, various health-related apps have been launched, including apps for sexually transmitted diseases. Until now, little has been known about how patients perceive the value of such apps.

Objective: To investigate patient's attitudes and awareness toward sexually transmitted disease–related apps in an outpatient sexually transmitted disease clinic setting.

Methods: A cross-sectional study was conducted at a dermatovenereological outpatient unit between April and July 2019. Patients completed a self-administered questionnaire on their perceptions of the popularity and usefulness of sexually transmitted disease–related apps. Descriptive analysis was performed with expression of categorical variables as frequencies and percentages. For continuous variables, the median, range, and interquartile range were indicated. Contingency tables and chi-square tests were used to investigate associations between sociodemographic data and items of the questionnaire.

Results: A total of 226 patients were surveyed (heterosexual: 137/193, 71.0%; homosexual: 44/193, 22.8%; bisexual: 12/193, 6.2%); 11.9% (27/225) had previously used health-related apps. Nearly half of the patients (97/214, 45.3%) specifically considered sexually transmitted disease–related apps useful, 47.8% (100/209) voted that they could supplement or support the consultation of a physician. Interestingly, only 35.1% (74/211) preferred a printed patient brochure on sexually transmitted diseases over downloading and using an app, but 64.0% (134/209) would download a sexually transmitted disease–related app recommended by their physician. General information regarding sexually transmitted diseases (93/167, 55.7%), evaluation of skin diseases based on photos or videos (78/167, 53.3%), information on the prevention of sexually transmitted diseases (76/167, 45.5%), mediation of nearby contact points or test sites (74/167, 44.3%), anonymous medical advice (69/167, 41.3%), and calculation of the risk of having a sexually transmitted disease (63/167, 37.3%) were rated as the most important features. Men were more likely than women to find sexually transmitted disease–related apps useful in general (P=.04; χ^2=6.28) and to pay for such apps (P=.01; χ^2=9.19). Patients aged <40 years would rather download an app recommended by their physician (P=.03; χ^2=7.23), whereas patients aged >40 years preferred reading a patient brochure on sexually transmitted diseases (P=.02; χ^2=8.14).

Conclusions: This study demonstrated high general interest in the use of sexually transmitted disease–related apps in this sample of dermatovenereological outpatients. In particular, young age and male sex were significantly associated with a positive perception, underlining the high potential of apps in the prevention and early recognition of sexually transmitted diseases in this group. Future studies are warranted to validate these findings in other populations.

(JMIR Mhealth Uhealth 2020;8(11):e16517) doi:10.2196/16517

KEYWORDS

sexually transmitted diseases; sexually transmitted infection; mobile phone apps; health apps

Introduction

Mobile phones and tablets are being increasingly integrated into the daily lives of many people worldwide. Mobile health apps have a high potential to improve the quality and coverage of care; increase access to health information, services, and skills; as well as promote changes in health behaviors [1]. However, there is little regulatory control over the accuracy and medical validity of the information provided in apps [2]. Furthermore, privacy, moral, and ethical concerns have been raised [3,4]. Data extracted from digital health apps could be used to record, survey, monitor, influence, and discipline users [5].

Sexually transmitted diseases (STDs) represent a common but preventable cause for morbidity and serious medical complications. In our clinical experience, patients of STD clinics are usually young and especially prone to the use of modern technologies. Digital revolution did not only change their way of social interaction, in fact, in our clinical experience, it changed the risks associated with their sexual behavior, especially among men having sex with men. Meeting sex partners through geosocial networking apps is common among men having sex with men [6]. The use of internet- and app-based services, such as Grindr or Tinder, to facilitate sexual partnering is a known risk factor for the transmission of STDs and human immunodeficiency virus (HIV) [7-10].

In response, many online dating apps have established standard notification methods to inform users of recent HIV or STD exposure [7,11]. Other functions of health apps in the context of STDs aim to ensure adherence to antiretroviral therapy medication in patients who are HIV-positive by reminding them of antiretroviral therapy intake, providing health care via video consultations (mobile telemedicine), or through self-test kits (home tests) for detection of STDs [12,13]. Other common features of STD-related apps include sexual health promotion and educational information on prevention of STDs [14].

Given the strong emergence of various STD-related health apps, it is crucial to comprehend their implication in daily life as well as to document user experience and perception of those tools from a patients' perspective. However, until now, little was known about how patients perceive the value of such apps. Hence, we aimed to examine how patients from an STD clinic perceive mobile health apps for STDs with a cross-sectional study.

Methods

Design and Procedure

A cross-sectional study including patients from the dermatovenereological outpatient clinic of the dermatology and allergy department of the university hospital of Munich was conducted between April and July 2019. This study was approved by the institutional review board of the university hospital (Ludwig Maximilian University Munich; approval number 19-336 KB). We closely adhered to the STROBE statement for cross-sectional studies for the reporting of this study [15,16].

Setting, Participants, and Sampling

The dermatovenereological unit of the university hospital of Munich is a public outpatient clinic in downtown Munich, Germany. Munich represents a wealthy city with comparably low crime and drug use as well as high levels of education [17]. The unit mainly focuses on the diagnosis, treatment, and follow-up care of patients with sexually transmitted infections. Furthermore, patients with nonvenereal genital diseases such as candidiasis or chronic inflammatory dermatoses represent an important number of cases. These 2 patient groups are different in many ways. Patients presenting with STDs are most commonly young men having sex with men. In contrast, patients suffering from inflammatory diseases of the anogenital area are frequently postmenopausal women. During the period when the survey was conducted (April to July 2019), 969 individuals (1241 visits) presented to the unit, with a median age of 40.3 years, of which, 73.1% were male and 26.9% were female. The top 3 diagnoses using International Statistical Classification of Diseases, Tenth Revision (ICD-10) codes were HIV (B24.0; n=319), condyloma acuminata (n=190), and other STDs (A64.0; n=75). In that period, there were 104 patients with nonvenereal diseases that included multiple distinct diagnoses according to ICD-10 (L90.0, L30.9). All German-speaking adult patients (16 years or older) presenting to the unit were asked to complete a 3-page paper questionnaire by a physician (LJ, ZF, TP). Participation was voluntary, and all participants gave verbal informed consent in German before completing the questionnaire. The individual questions were not linked (ie, individual questions could be skipped to continue with the questionnaire). Refusals were not documented, and no incentives were provided. The study population was a convenience sample.

Survey

As no validated survey tools for the objective of our study existed, the 3-page questionnaire was developed de novo based on a literature review and dermatovenereological expert

consulting, and included questions on STD-related apps and basic demographic information (age; gender; place of residence—town, suburb, or rural area; and education).

In a multiple-choice question format, patients were asked about the reason for presenting to the unit on the day of the assessment, their sexual orientation, and their sexual risk behavior. Other studies [18] that exclusively focus on STDs in a high-risk population have used criteria such as the number of condomless anal intercourse encounters as risk criteria for acquiring an STD. By considering the number of patients with noninfectious diseases of the anogenital area, we estimated that the individual risk level for STDs for those in our study population was moderate. We asked participants if they had changed partners or had a sexual disease recently. Thus, the sexual risk behavior was operationalized by the number of sexual partners in the previous 6 months. This value was decided by our clinical experience in this specific population. Other questions addressed previous use of health-, and specifically, STD-related apps, perceptions of the relevance of those apps, as well as concerns regarding digital security. These questions were dichotomous; however, patients could also state that they were unsure. The full questionnaire is available in Multimedia Appendix 1. It was validated by 10 healthy volunteers (4 male, 6 female) for clarity and comprehensibility, and 5 physicians assessed the questionnaire for clinical validity referring to an external benchmark study [12] that was similar in content to our investigation. Based on their feedback, the questionnaire was revised to its final form. Completed questionnaires were sequentially numbered for data entry purposes but were not linked to any identifying patient information to assure irreversible anonymity.

Data Analysis

Categorical variables were expressed as frequencies and percentages and were compared using 2-sided chi-square tests. $P<.05$ was considered statistically significant. For continuous variables, median, range, and interquartile range were reported. Statistical analyses were conducted with SPSS statistical software (version 25; IBM Corp).

Results

Characteristics of the Study Population

A total of 226 patients were included. Sociodemographic characteristics of the population are shown in Table 1.

The majority (75.2%, 170/226) presented with genital discomfort, which persisted for more than 6 weeks in 69.2% of patients (117/169). Additionally, most patients had already consulted a physician in person for their symptoms (124/162, 76.5%). Further reasons for presentation were examination for STDs (72/226, 31.9%), HIV treatment or consultation (20/226, 8.8%), skin cancer screening (16/226, 7.1%), or other (54/226, 23.9%).

The majority of patients had 0-2 different sex partners in the previous 6 months (126/176, 71.6%), and 28.4% (50/176) had more than 2 sex partners in the previous 6 months. Overall, 21.5% of patients had already been diagnosed with an STD prior to presentation (44/205). Regarding sexual orientation, patients were heterosexual (137/193, 71.0%), homosexual (44/193, 22.8%), or bisexual (12/193, 6.2%). Interestingly, 28.0% (58/207) had already made use of an app to arrange sexual encounters, underlining the moderate to high-risk sexual behavior of the study population (Table 1).

Table 1. Overview of the characteristics of the included patients in the cross-sectional study (N=226).

Characteristic	Value
Sex (n=210), n (%)	
Female	49 (23.3)
Male	161 (76.7)
Age (in years; n=209)	
median (IQR)	37 (27-52)
range	16-83
mean (SD)	40.2 (15.2)
Education (n=206), n (%)	
University degree	67 (32.5)
General higher education entrance qualification	57 (27.7)
Secondary school leaving certificate	47 (22.8)
Lower secondary school leaving certificate	28 (13.6)
Other degree	1 (0.5)
No degree	6 (2.9)
Residence (n=210), n (%)	
City or suburb	165 (78.6)
Rural area	45 (21.4)
Reason for appointment (n=226, multiple answers possible), n (%)	
Examination for STD	72 (31.9)
Skin cancer screening	16 (7.1)
HIV treatment or consultation	20 (8.8)
Other reasons (eg, follow-up care)	54 (23.9)
Genital discomfort	170 (75.2)
Genital discomfort persisting for more than 6 weeks (n=169)	117 (69.2)
Previous consultation of a physician due to genital discomfort (n=162)	124 (76.5)
Previous diagnosis of an STD (≤6 months) (n=205), n (%)	
Yes	44 (21.5)
No	161 (78.5)
Sexual orientation (n=193), n (%)	
Heterosexual	137 (71.0)
Homosexual	44 (22.8)
Bisexual	12 (6.2)
Number of sex partners (≤6 months) (n=176), n (%)	
0-2 sex partners	126 (71.6)
>2 sex partners	50 (28.4)
Previous use of an app to arrange sexual encounters (n=207), n (%)	
Yes	58 (28.0)
No	149 (72.0)

Previous Experience With Health Apps

In total, 34.1% (76/223) of the patients owned both a smartphone and a tablet, 59.6% (133/223) were owners of a smartphone only, and 2.2% (5/223) were owners of a tablet device only. Only 4.0% (9/223) had none of these devices. Most of patients (184/226, 81.4%) had already searched the internet prior to this survey in order to obtain health-related information, and

additionally, 72.0% (152/211) had searched the internet before to acquire information about the specific complaints for which they were visiting the clinic.

When asked about previous experience with health-related apps, 11.9% (27/225) had previously made use of such apps, whereas the overwhelming majority (194/225, 86.7%) denied or were unsure (3/225, 1.3%). Apps that had already been used by the patients had been provided by health insurance (*DAK-Gesundheit*, *Allgemeine Ortskrankenkasse*, *Barmer Ersatzkasse*, *Techniker Krankenkasse*) or had been health-tracking apps offered by Apple, Android, Runtastic (fitness-tracking app), or Garmin.

Most patients (110/166, 66.3%) rated scientifically reliable information as the most important feature for health-related apps. For 52.4% (87/166), credibility of the provider as well as data security were important, followed by user convenience (86/166, 51.8%) and continuous availability (78/166, 47.0%); 41.0% (68/166) and 27.1% (45/166) considered a low price and an attractive layout, respectively, as critical.

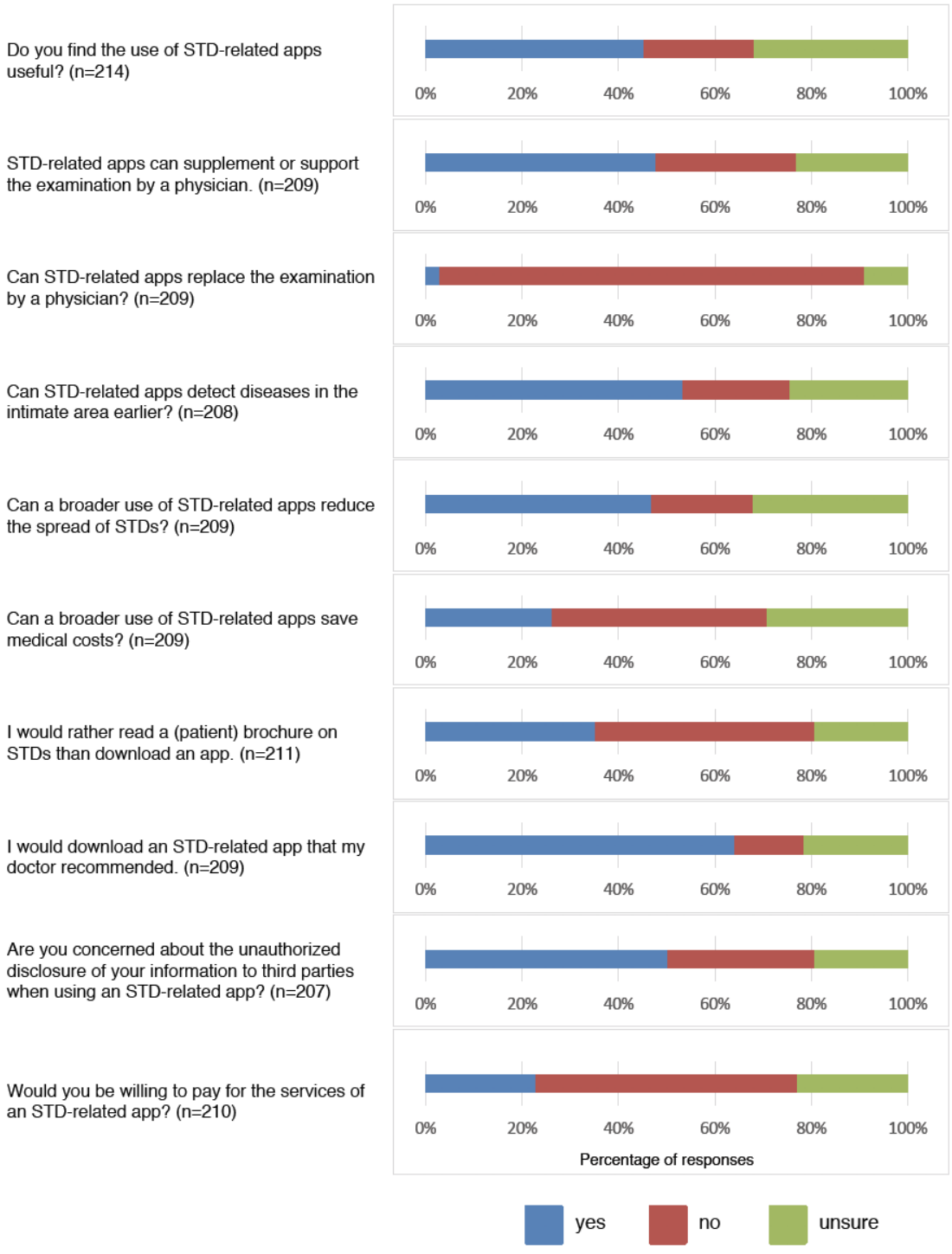
Attitude Toward STD-Related Apps

Nearly half of the patients (97/214, 45.3%) considered apps for patients with complaints in the intimate area or venereal diseases to be useful; 31.8% (68/214) and 22.9% (49/214) were unsure or thought that such apps were not useful for patients, respectively; and 47.8% (100/209) voted that STD-related apps could supplement or support the consultation of a physician, while 23.0% (48/209) were unsure. The vast majority did not believe that apps can replace the consultation by a physician (cannot be replaced: 184/209, 88.0%; unsure: 19/209, 9.1%). More than half of the patients (111/208, 53.4%) thought that apps can contribute to detecting diseases in the intimate area earlier. In addition, 46.9% of patients (98/209) were convinced that broad usage may reduce the spread of STDs (unsure: 67/209, 32.1%; cannot reduce the spread: 44/209, 21.1%). Furthermore, less than half of the patients did not agree with the statement that the broad use of such apps can save medical costs (disagree: 493/209, 4.5%; unsure: 61/209, 29.2%).

Interestingly, only 35.1% of patients (74/211) preferred a printed patient brochure on STDs over downloading and using an app. Half of the patients (104/207, 50.2%) had concerns about the unauthorized disclosure of information to third parties, while 30.4% (63/207) did not have concerns and 19.3% (40/207) were unsure. However, 64.0% of the patients (134/209) would download an STD-related app recommended by their physician, and only 22.9% (48/210) were willing to pay for services provided by an STD-related app (Figure 1).

We also asked the patients which app features they considered important for patients with complaints in the intimate area. The majority reported general information regarding STDs to be important (93/167, 55.7%), followed by the evaluation of skin diseases based on photos or videos (78/167, 53.3%), information on the prevention of STDs (45.5%76/167), mediation of nearby contact points or test sites (74/167, 44.3%), anonymous medical advice (69/167, 41.3%), and calculation of the risk of having an STD (63/167, 37.3%). Less than one-third of the patients found treatment plans with reminders for medication (46/167, 26.9%), mailing of test kits to be used at home (43/167, 25.7%), and verification of STD status of sexual partners or safe dating (24/167, 14.4%) to be relevant.

Figure 1. Patients' answers regarding their attitudes toward STD apps.



Association Between Sociodemographic Data And Attitude Toward STD-Related Apps

Sex (Men Versus Women)

In comparison to women, men were more likely to find STD-related apps useful in general (P=.04; χ^2=6.28). However, in terms of possible features, men were less frequently interested than women in apps on information about prevention of STDs (P=.02; χ^2=5.82). In our sample, more men than women agreed that STD-related apps can detect STDs earlier (P<.001; χ^2=15.72) and that medical costs can be saved (P=.006; χ^2=10.37). In addition, more men than women preferred reading a brochure on STDs over downloading an app (P=.006; χ^2=10.18). There was also an association between the willingness to pay for the services provided by an app and male sex (P=.01; χ^2=9.19)—men were more willing to pay than women were (Table 2).

Table 2. Association between sociodemographic data and attitude toward STD-related apps.

Item of the questionnaire	*P* value			
	Sex (male vs female)	Residence (city/suburb vs rural)	Age (<40 years vs ≥40 years)	Sexual risk behavior (0-2 vs >2 partners)
Generally useful?	.04	.08	.08	.04
Supplement or supportive?	.51	.13	.25	.63
Replace examination?	.45	.55	.27	.09
Earlier detection?	.001	.027	.001	.02
Reduce spread of STDs?	.07	.13	.11	.02
Save medical costs?	.006	.03	.009	.06
Prefer brochure?	.006	.01	.02	.16
Download recommended app?	.16	.48	.027	.72
Disclosure of personalized data?	.24	.44	.12	.12
Willing to pay?	.01	.04	.052	.29

Residence (City/Suburb Versus Rural Area)

Patients living in a city or suburb versus those living in rural areas agreed more frequently to the statements—STD-related apps can detect diseases in the intimate area earlier (P=.03; χ^2=7.25) and apps can contribute to saving medical costs (P=.03; χ^2=6.75). Furthermore, patients living in a city or suburb, compared to participants from rural areas, were more willing to pay for the services offered in such apps (P=.04; χ^2=6.73).

Age (<40 Years Versus ≥40 Years)

Patients <40 years of age considered an attractive layout (P=.04; χ^2=4.25), evaluation of skin diseases based on photos or videos (P=.004; χ^2=8.498), and mailing of test kits for home usage (P=.002; χ^2=9.77) as important features for STD apps more frequently than those >40 years of age. Additionally, patients <40 years were more likely to agree that diseases may be detected in a timely manner (P< 0.001; χ^2=18.82) and that medical costs can be saved through the usage of STD-related apps (P=.009; χ^2=9.40) than those >40 years of age. Furthermore, patients <40 years of age would rather download an app recommended by their physician instead of reading a printed brochure (P=.03; χ^2=7.23), whereas patients aged >40 years preferred reading a patient brochure on STDs instead of downloading an app (P=.02; χ^2=8.14).

Sexual Risk Behavior (0-2 Partners Versus >2 Partners in the Previous 6 Months)

Regarding preferred features of STD-related apps, evaluation of skin diseases based on photos or videos was chosen more frequently by patients with 0-2 sexual partners in the previous 6 months than it was chosen by those with >2 partners (P=.01; χ^2=6.19). In addition, patients with 0-2 sexual partners in the previous 6 months were more interested in home test kit mailing for STD self-tests (P=.04; χ^2=4.38) than those with >2 partners in the previous 6 months were. Patients with >2 partners were more interested in safe sex partner dating or verification of the STD status of the partner than those with 0-2 partners were (P=.003; χ^2=8.86).

Other Factors

Patients without an STD in the previous 6 months indicated more frequently that self-test kits for STDs (P=.005; χ^2=8.05) and individual STD risk assessment (P=.008; χ^2=7.05) were important features of STD apps compared to those with a recent STD. Patients who did not use apps to arrange sexual encounters chose the feature evaluation of skin diseases based on photos or videos (P=.001; χ^2=11.797) more frequently than those regularly using apps did. Mailing of test kits (P<.001; χ^2=20.64) was found to be crucial for those who had already made use of dating apps in the past than for those who had not used any apps.

Discussion

In recent years, scientists have evaluated many apps in the field of sexual health [14,19]. We believe it is necessary to ask the target groups of STD-related apps about all aspects of those digital services directly, to explore the patient perspective. The overall goal of this study was to explore the patient perspective of STD-related apps, and thus, to guide the further development of apps based on these identified patient-centered preferences.

We interviewed 226 patients, who presented to an outpatient STD clinic for various reasons, about their perception of STD-related apps. The study population was heterogeneous and included both women and men, with men being predominant. In our clinical experience, the proportion of men having sex with men was low compared to those of other STD clinics. This may be explained by the fact that this STD clinic belongs to the department of dermatology where many patients are treated with chronic inflammatory conditions or other primary nonvenereal diseases of the genital area. Nevertheless, 28.4% of the patients (126/176) reported more than 2 sexual partners in the previous 6 months, and 28.0% (58/207) indicated that they had already used an app at least once to make sexual

contacts. From these data, we conclude that the sexual behavior of the sample presented in this study is most likely to correspond to moderate to high risk for STDs.

To our knowledge, this is the first study to include an analog survey of patients seeking medical advice from a physician on their perception of the many aspects of digital health services regarding sexual health. In contrast to the focus of other studies [20,21], we did not focus exclusively on specific high-risk groups such as men having sex with men, individuals who are HIV-positive, or users of geosocial sexual networking apps but interviewed the patients regardless of their sexual orientation, HIV status, or internet expertise to gain a broad opinion of STD-related apps. We believe that this approach is necessary for 2 reasons. First, many existing STD apps are aimed at a broad usership. Second, in our experience, patients with genital discomfort usually do not know for themselves whether their complaints are due to an STD or some other nonsexually transmitted disease. In this study, we intended to encompass the perspective of these distinct patient groups.

Despite the fact that most of the patients surveyed in this study had no prior experience with health apps, the majority used the internet to get health information in general: 72% of the study population (152/211) affirmed that they had searched the internet for information about the specific complaint that they presented at the STD clinic and, additionally, 64.0% (134/209) indicated that they would download an STD app recommended by their physician. These findings demonstrate a great general interest of STD patients in acquiring digital health information before and after visiting a doctor. It also indicates that even if the vast majority did not believe that apps could replace a consultation by a physician, they would generally appreciate additional information.

There might be a group of patients with similar complaints who had looked up information on the internet, was satisfied, and thus never presented to the clinic, and therefore, was not represented in the study. However, surveys of physicians who offered app-based care for STD patients showed referral rates to clinics of almost 100% [12]. This and other studies [14,19] that evaluated a range of STD-related apps support the statements of our study population that mobile medical apps in the field of STDs do not replace visits to clinics and form an argument against the existence of this second unrepresented group.

Interestingly, despite the low experience of our study population with health apps, the majority preferred to obtain health information delivered via app over reading a patient brochure in contrast to the findings of other studies [22]. Thus, the results highlight the great potential for digital health app providers in targeting multiple patient groups with complaints in the genital area.

When questioned about important aspects of digital health apps, the most valued aspects were scientifically reliable information, followed by data security and a trustworthy provider. These items highlight the strong need for reliability and confidentiality. Of all functions, general information and the evaluation of skin diseases based on photos or videos were the features most frequently desired. We estimate that these features could be useful for both patients with STDs and patients with nonvenereal genital complaints. In contrast, functions to find partners for safe sex, home-based tests, and treatment plans with reminders of medication intake (eg, for antiretroviral therapy) were chosen less frequently. This could be due to the fact that high-risk groups such as men having sex with men or patients who are HIV-positive were relatively underrepresented in this population [20]. In particular, the choice of apps for safe sex partner dating and verification of the STD status of the partner was very significantly associated with the number of sexual contacts within the previous 6 months, highlighting that individual sexual behavior is closely related to desired target functions of the apps. In contrast, patients with fewer sexual partners within the previous 6 months were more interested in an evaluation and recognition of genital diseases, and especially, in home-based tests. These results underline the great interest of patients in an anonymous diagnosis or consultation without the necessity of a direct contact with a physician.

We are aware that this study has several limitations. The sample comprised 226 patients presenting to the venereological outpatient unit. First, this sample size is relatively small, and second, it was not sampled in a random fashion but depending on the availability of patients. The questionnaire has not been validated and was developed de novo. Thus, the questionnaire may lack validity and the results presented here may not be fully generalizable to other populations or other STD clinics and are at risk of sampling bias. Although the survey was anonymized, patients were likely known to the treating physicians, which may have altered the response in either direction, potentially biasing the results. The items of the questionnaire focused primarily on what patients might do outside of a clinic and did not include clinic-specific aspects such as communication with staff. Furthermore, this sample was a selection of patients physically presenting to the unit. Thus, the value and evaluation of STD-related apps among patients who did not present to seek STD consultation and treatment remains unclear. The response rate on the individual questions varied but was generally high. The question with the lowest response rate (77.9%) was the one about the number of sexual partners. This is consistent with other studies that investigated sexual risk behavior [20]. It could be due to the fact that participants perceive answering intimate questions as uncomfortable or, despite being assured anonymity, they are not sure about the confidential treatment of their data [18].

In our study population, men were generally more open to the use of STD-related apps than female patients. In comparison to female participants, men rated STD-related apps as more useful, cost-effective, and more reliable in early detection of STDs. Furthermore, they were more likely than females were to be willing to spend money for the services of an STD-related app. This may be explained by the gender disparity of our study population in terms of distribution of STDs and nonvenereal genital diseases. In our STD department, we observe predominantly male patients with STDs. Furthermore, we tend to have more female patients presenting with nonvenereal genital inflammatory diseases. This perception is consistent with epidemiologic data on the sex distribution of STDs in Germany [23]. Gender disparity in the distribution of STDs could be a

reason for the higher interest of male participants in STD-related apps in general, but could it possibly explain why, in this study, men showed significantly less interest in apps with information on prevention of STDs than women? Higher sexual risk behavior might be associated with a lack of interest in prevention of STDs. We believe these data underline the hypothesis that women and men as well as specific risk groups should be addressed individually with tailored information in order to achieve a higher engagement in adopting preventive behavior. Further studies are needed in order to improve future STD prevention and information efforts and to validate our findings in other populations.

Conflicts of Interest

TJB is the owner of Smart Health Heidelberg GmbH which provides the STD-triage app Intimarzt. LF has consulted and given lectures for Galderma, Janssen, Leo Pharma, Eli Lilly, Abbvie, Amgen and Novartis.

Multimedia Appendix 1
Original questionnaire.
[PDF File (Adobe PDF File), 221 KB - mhealth_v8i11e16517_app1.pdf]

References

1. mHealth: New horizons for health through mobile technologies. World Health Organization. 2011. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf? [accessed 2020-10-20]
2. Kessel KA, Vogel MM, Kessel C, Bier H, Biedermann T, Friess H, et al. Mobile health in oncology: a patient survey about app-assisted cancer care. JMIR Mhealth Uhealth 2017 Jun 14;5(6):e81 [FREE Full text] [doi: 10.2196/mhealth.7689] [Medline: 28615159]
3. Huckvale K, Prieto JT, Tilney M, Benghozi P, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015;13:214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]
4. Charani E, Castro-Sánchez E, Moore LSP, Holmes A. Do smartphone applications in healthcare require a governance and legal framework? It depends on the application!. BMC Med 2014;12:29 [FREE Full text] [doi: 10.1186/1741-7015-12-29] [Medline: 24524344]
5. Lupton D. M-health and health promotion: the digital cyborg and surveillance society. Soc Theory Health 2012 Jun 27;10(3):229-244. [doi: 10.1057/sth.2012.6]
6. Goedel WC, Duncan DT. Geosocial-networking app usage patterns of gay, bisexual, and other men who have sex with men: survey among users of Grindr, a mobile dating app. JMIR Public Health Surveill 2015;1(1):e4 [FREE Full text] [doi: 10.2196/publichealth.4353] [Medline: 27227127]
7. Contesse MG, Fredericksen RJ, Wohlfeiler D, Hecht J, Kachur R, Strona FV, et al. Attitudes about the use of geosocial networking applications for HIV/STD partner notification: a qualitative study. AIDS Educ Prev 2019 Jun;31(3):273-285 [FREE Full text] [doi: 10.1521/aeap.2019.31.3.273] [Medline: 31145001]
8. Landovitz RJ, Tseng C, Weissman M, Haymer M, Mendenhall B, Rogers K, et al. Epidemiology, sexual risk behavior, and HIV prevention practices of men who have sex with men using GRINDR in Los Angeles, California. J Urban Health 2013 Aug;90(4):729-739 [FREE Full text] [doi: 10.1007/s11524-012-9766-7] [Medline: 22983721]
9. Rosser BRS, Oakes JM, Horvath KJ, Konstan JA, Danilenko GP, Peterson JL. HIV sexual risk behavior by men who use the Internet to seek sex with men: results of the Men's INTernet Sex Study-II (MINTS-II). AIDS Behav 2009 Jun;13(3):488-498 [FREE Full text] [doi: 10.1007/s10461-009-9524-3] [Medline: 19205866]
10. Wang H, Zhang L, Zhou Y, Wang K, Zhang X, Wu J, et al. The use of geosocial networking smartphone applications and the risk of sexually transmitted infections among men who have sex with men: a systematic review and meta-analysis. BMC Public Health 2018 Oct 16;18(1):1178 [FREE Full text] [doi: 10.1186/s12889-018-6092-3] [Medline: 30326887]
11. Mobley V, Cope A, Dzialowy N, Maxwell J, Foust E, Samoff E. A comparison of syphilis partner notification outcomes by reported use of internet-based apps to meet sex partners in North Carolina, 2013-2016. Sex Transm Dis 2018 Dec;45(12):823-828. [doi: 10.1097/OLQ.0000000000000884] [Medline: 29944644]
12. Cao B, Zhao P, Bien-Gund C, Tang W, Ong JJ, Fitzpatrick T, et al. The web-based physician is ready to see you: a nationwide cross-sectional survey of physicians using a mobile medical app to evaluate patients with sexually transmitted diseases in China. JMIR Mhealth Uhealth 2018 Oct 30;6(10):e10531 [FREE Full text] [doi: 10.2196/10531] [Medline: 30377148]
13. Morano JP, Clauson K, Zhou Z, Escobar-Viera CG, Lieb S, Chen IK, et al. Attitudes, beliefs, and willingness toward the use of mhealth tools for medication adherence in the Florida mhealth adherence project for people living with HIV (FL-mAPP): pilot questionnaire study. JMIR Mhealth Uhealth 2019 Jul 03;7(7):e12900 [FREE Full text] [doi: 10.2196/12900] [Medline: 31271150]
14. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J Med Internet Res 2013;15(1):e1 [FREE Full text] [doi: 10.2196/jmir.2301] [Medline: 23291245]

15. Vandenbroucke JP, von EE, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med 2007 Oct 16;4(10):e297 [FREE Full text] [doi: 10.1371/journal.pmed.0040297] [Medline: 17941715]
16. von EE, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Int J Surg 2014 Dec;12(12):1495-1499 [FREE Full text] [doi: 10.1016/j.ijsu.2014.07.013] [Medline: 25046131]
17. Kommunaler Finanzreport 2019. Bertelsmann Stiftung. URL: https://www.bertelsmann-stiftung.de/fileadmin/files/Projekte/Kommunale_Finanzen/Finanzreport-2019-gesamt.pdf [accessed 2020-10-20]
18. Mor Z, Shohat T, Goor Y, Dan M. Risk behaviors and sexually transmitted diseases in gay and heterosexual men attending an STD clinic in Tel Aviv, Israel: a cross-sectional study. Isr Med Assoc J 2012 Mar;14(3):147-151. [Medline: 22675852]
19. Gibbs J, Gkatzidou V, Tickle L, Manning SR, Tilakkumar T, Hone K, et al. 'Can you recommend any good STI apps?' a review of content, accuracy and comprehensiveness of current mobile medical applications for STIs and related genital infections. Sex Transm Infect 2017 Jun;93(4):234-235. [doi: 10.1136/sextrans-2016-052690] [Medline: 27884965]
20. Werner RN, Gaskins M, Ahrens J, Jessen H, Kutscha F, Mosdzen R, et al. Knowledge and use of HIV pre-exposure prophylaxis among men who have sex with men in Berlin - a multicentre, cross-sectional survey. PLoS One 2018;13(9):e0204067 [FREE Full text] [doi: 10.1371/journal.pone.0204067] [Medline: 30212547]
21. Sun CJ, Stowers J, Miller C, Bachmann LH, Rhodes SD. Acceptability and feasibility of using established geosocial and sexual networking mobile applications to promote HIV and STD testing among men who have sex with men. AIDS Behav 2015 Mar;19(3):543-552. [doi: 10.1007/s10461-014-0942-5] [Medline: 25381563]
22. Steeb T, Wessely A, Niesert A, Ruzicka T, von Braunmühl T, Berking C, et al. Patient attitude towards videodermatoscopy for the detection of skin cancer: a cross-sectional study. Oncol Res Treat 2019;42(6):319-325. [doi: 10.1159/000499630] [Medline: 30995670]
23. Bremer V, Dudareva-Vizule S, Buder S, an der Heiden M, Jansen K. Sexuell übertragbare Infektionen in Deutschland. Bundesgesundheitsbl 2017 Jul 24;60(9):948-957. [doi: 10.1007/s00103-017-2590-1]

Abbreviations

HIV: human immunodeficiency virus
ICD-10: International Statistical Classification of Diseases, Tenth Revision
STD: sexually transmitted disease
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by G Eysenbach; submitted 06.10.19; peer-reviewed by G Hobson, A Dinkel, R Cook, T Peng; comments to author 18.05.20; revised version received 08.08.20; accepted 04.09.20; published 10.11.20.

Please cite as:
Jakob L, Steeb T, Fiocco Z, Pumnea T, Jakob SN, Wessely A, Rothenberger CC, Brinker TJ, French LE, Berking C, Heppt MV
Patient Perception of Mobile Phone Apps for the Care and Prevention of Sexually Transmitted Diseases: Cross-Sectional Study
JMIR Mhealth Uhealth 2020;8(11):e16517
URL: https://mhealth.jmir.org/2020/11/e16517
doi:10.2196/16517
PMID:33170133

Review

Standalone Smartphone Cognitive Behavioral Therapy–Based Ecological Momentary Interventions to Increase Mental Health: Narrative Review

Marta Anna Marciniak[1], MSc; Lilly Shanahan[2], PhD; Judith Rohde[1], MD; Ava Schulz[1], PhD; Carolin Wackerhagen[3], PhD; Dorota Kobylińska[4], PhD; Oliver Tuescher[5], MD; Harald Binder[6], PhD; Henrik Walter[3], MD; Raffael Kalisch[5], PhD; Birgit Kleim[1], PhD

[1]University of Zurich, Psychiatric University Hospital, Zurich, Switzerland

[2]Jacobs Centre of Productive Youth Development, University of Zurich, Zurich, Switzerland

[3]Charite Berlin Universitätsmedizin, Berlin, Germany

[4]University of Warsaw, Warsaw, Poland

[5]Gutenberg University Medical Center, Mainz, Germany

[6]Institute for Medical Biometry and Statistics, Faculty of Medicine, University of Freiburg, Freiburg, Germany

Corresponding Author:
Birgit Kleim, PhD
University of Zurich
Psychiatric University Hospital
Lenggstrasse 31
Zurich, 8032
Switzerland
Phone: 41 44 384 23 51
Email: birgit.kleim@uzh.ch

Abstract

Background: A growing number of psychological interventions are delivered via smartphones with the aim of increasing the efficacy and effectiveness of these treatments and providing scalable access to interventions for improving mental health. Most of the scientifically tested apps are based on cognitive behavioral therapy (CBT) principles, which are considered the gold standard for the treatment of most mental health problems.

Objective: This review investigates standalone smartphone-based ecological momentary interventions (EMIs) built on principles derived from CBT that aim to improve mental health.

Methods: We searched the MEDLINE, PsycINFO, EMBASE, and PubMed databases for peer-reviewed studies published between January 1, 2007, and January 15, 2020. We included studies focusing on standalone app-based approaches to improve mental health and their feasibility, efficacy, or effectiveness. Both within- and between-group designs and studies with both healthy and clinical samples were included. Blended interventions, for example, app-based treatments in combination with psychotherapy, were not included. Selected studies were evaluated in terms of their design, that is, choice of the control condition, sample characteristics, EMI content, EMI delivery characteristics, feasibility, efficacy, and effectiveness. The latter was defined in terms of improvement in the primary outcomes used in the studies.

Results: A total of 26 studies were selected. The results show that EMIs based on CBT principles can be successfully delivered, significantly increase well-being among users, and reduce mental health symptoms. Standalone EMIs were rated as helpful (mean 70.8%, SD 15.3; n=4 studies) and satisfying for users (mean 72.6%, SD 17.2; n=7 studies).

Conclusions: Study quality was heterogeneous, and feasibility was often not reported in the reviewed studies, thus limiting the conclusions that can be drawn from the existing data. Together, the studies show that EMIs may help increase mental health and thus support individuals in their daily lives. Such EMIs provide readily available, scalable, and evidence-based mental health support. These characteristics appear crucial in the context of a global crisis such as the COVID-19 pandemic but may also help reduce personal and economic costs of mental health impairment beyond this situation or in the context of potential future pandemics.

(JMIR Mhealth Uhealth 2020;8(11):e19836) doi:10.2196/19836

KEYWORDS

mHealth; mobile app; ecological momentary intervention; EMI; cognitive behavioral therapy; CBT; COVID-19; mobile phone; smartphone

Introduction

Prevalence of Psychiatric Disorders

Heightened prevalence of psychiatric disorders is one of the largest challenges for modern health care. In 2001, the World Health Organization (WHO) estimated that 1 in every 4 individuals worldwide was affected by one or more mental or neurological health problems during their lives [1]. Prospective longitudinal studies suggest that the vast majority of people will have a mental health disorder at some point in their lives [2,3]. In 2017, >970 million people were diagnosed with mental or substance use disorders [4]. These numbers likely represent a gross underestimate because of underreporting for fear of stigma or limited mental health literacy, limited access to therapy services, high costs of treatment, and additional reasons, such as recent conditions caused by the COVID-19 outbreak.

The global COVID-19 pandemic is likely to further increase the need for mental health care interventions and new ways of implementing them. Global pandemics cause high levels of stress and may lead to mental health problems, such as depression, anger, anxiety disorders, and posttraumatic stress disorder, as well as an increase in smoking and alcohol consumption [5,6]. Current measures of physical distancing and quarantine aimed at curbing the spread of COVID-19 likely have additional detrimental psychological side effects, including loneliness [7]. At the same time, people are considerably less likely to receive professional face-to-face psychological help to overcome their fears, lowered mood, and other mental health problems [8].

Given the low detection rates of mental health problems, the WHO estimates that 76% to 85% of people in low- and middle-income countries and 35% to 50% of people in high-income countries receive no treatment for their mental health disorder [9]. After the pandemic is over, these numbers will be even higher. Moreover, those who seek treatment often do so only years into the mental disease, meaning that comorbid disorders and difficulties in many domains of life, including family and work, have already developed [10].

Ecological Momentary Interventions

The rapid growth in the use of smartphones has created a new branch of medicine—mobile health (mHealth). Currently, the most popular solutions in mHealth are mobile apps because they are easy to use and are widely available [11]. Such apps typically use ecological momentary intervention (EMI) to deliver treatments provided to people in their everyday lives [12]. This approach captures and modifies specific moment-to-moment situations that emerge in the real world rather than targeting problematic thoughts, emotions, and behaviors within therapy sessions or in the hospital [13,14].

Cognitive Behavioral Therapy

Most of the scientifically tested apps are based on cognitive behavioral therapy (CBT) principles. CBT was first created and established by Beck [15] and Ellis [16] and is based on the theory that maladaptive cognitions, such as general beliefs and automatic thoughts about the self and the world, contribute to the maintenance of emotional distress and behavioral problems. Accordingly, CBT specifically targets these maladaptive cognitions and behaviors. CBT is one of the most extensively used and researched form of psychotherapy [17] and is considered the gold standard for the treatment of many mental health problems [18]. Owing to their strong empirical foundation and clear structure, CBT-based interventions are well suited for application in mHealth. Here, we examine mobile apps that have been designed using a *rational app design*. With this definition, we refer to apps that are based on a CBT rationale and implement established and empirically validated CBT tools and techniques.

Objectives

Several reviews and meta-analyses reported the efficacy of EMIs for several mental health problems, including depression, anxiety, perceived stress, and eating disorders [12,19-23], but they were not focused solely on CBT-based interventions. Other reviews have focused on or have not excluded blended treatments such as face-to-face psychotherapy in combination with an app [24,25]. However, to the best of our knowledge, there are currently no reviews on the efficacy, effectiveness, and feasibility of standalone EMIs delivered via mobile apps and following a rational app design, in both healthy and clinical populations. However, the rapid development of this type of psychological support and recent pandemic-quarantine circumstances have created an urgent need for a comprehensive summary of the studies published so far.

Methods

Search Strategy

To build a comprehensive overview of the existing literature, 4 search terms were used to identify articles: (1) mental health, (2) smartphones, (3) CBT, and (4) ecological momentary interventions (see Multimedia Appendix 1 for the complete search strings) in the MEDLINE, PsycINFO, EMBASE, and PubMed databases on January 16, 2020.

Inclusion Criteria

The following criteria were used to select the studies: (1) peer-reviewed publications; (2) written in English; (3) published between January 1, 2007, and January 15, 2020; (4) standalone treatments (blended interventions were excluded); (5) explicitly aiming to increase mental health; and (6) focusing on feasibility, efficacy, and/or effectiveness of EMI. Both within- and

between-group designs and studies with both healthy and clinical samples were included.

The time range was decided based on the launch date of the first App Store in 2007. This means that the review proposes an overview of presumably all the papers published since the development of the field started.

Quality Assessment

For selected studies, we assessed the methodological risk of bias of included studies in accordance with the Cochrane Handbook [26] and the guidelines of the Cochrane Consumers and Communication Review Group [27], which recommends the explicit reporting of the following individual elements for the studies: random sequence generation, allocation sequence concealment, blinding (participants and personnel), blinding (outcome assessment), completeness of outcome data, selective outcome reporting, and other sources of bias. Overall, from the 14 randomized controlled trials (RCTs) included, 8 were rated as low risk and 6 as medium risk. From the remaining 12 non-RCT studies, 11 were rated as medium risk and 1 as high risk. The inter-rater reliability was 96%. Owing to the small number of studies fulfilling the inclusion criteria, we decided to keep the study with high risk in the review.

Data Extraction

Data extraction included (1) study design; (2) choice of the control condition; (3) sample characteristics in terms of sample size and justification of the size in power analysis, gender ratio, age, and mental health problems; (4) content of EMI in terms of specific CBT strategies; (5) EMI delivery characteristics in terms of mode, duration, and frequency; (6) feasibility in terms of acceptance, satisfaction, and helpfulness rates; (6) efficacy or effectiveness in terms of improvement in primary outcomes; and (7) outcome measures used in the studies.

Analysis

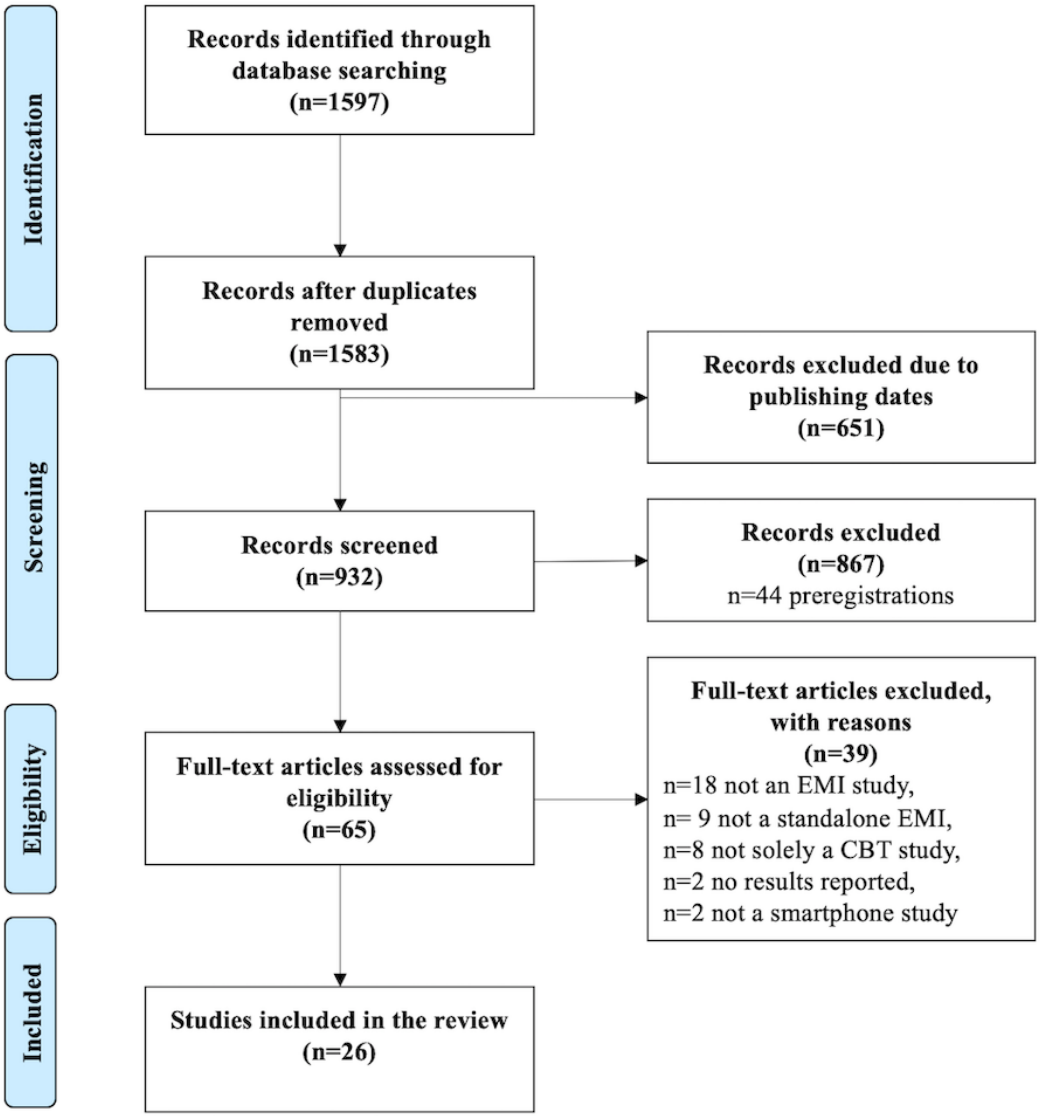
Owing to the heterogeneity of the designs and objectives of the included studies, quantitative analysis was possible only in a few cases, for instance, for sample characteristics, delivery characteristics, and feasibility measures. However, in many aspects, such as effectiveness and effectiveness measures, quantitative analysis was not possible and, instead, a narrative approach for the summary of the qualitative findings was chosen.

Results

Included Studies

Inclusion criteria resulted in the identification of 26 articles (see Figure 1 for the complete literature flow chart). Of those, 3 [28-30] described the outcomes of 1 trial. For this particular trial, we only included data from the latest publication to avoid bias or duplication. Thus, our results are based on n=24 studies.

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



Study Design and Control Conditions

A variety of study designs were employed to investigate the effects of EMIs: 2-, 3-, or 4-armed RCTs (12 studies [30-41]) as well as a nonrandomized trial with a control group (1 study [42]) for between-group comparisons or case studies and 1-group-only studies (11 studies [43-53]) for within-group or intraindividual comparisons (Table 1). Control conditions were selected based on the research questions of a given study. The control conditions in the studies reviewed here include (1) another app with different content (8 studies), (2) a waiting list (4 studies), (3) ecological momentary assessment (EMA)—participants were involved in self-monitoring only (2 studies), (4) encouragement to visit helpful websites or call hotlines (2 studies), (5) computer-based treatment (the same content of the intervention but delivered via internet browser; 1 study), (6) real-life training provided by a qualified person (eg, meditation training; 1 study), and (7) treatment as usual (1 study).

Table 1. Summary of the study design and control conditions.

Study	Names of apps	RCT[a]	Number of arms	Control condition	Type of control condition
Arevian et al, 2018 [43]	B-RESILIENT app	—[b]	—	—	—
Bakker et al, 2018 [32]	Mood Prism, Mood Mission, Mood Kit	✓[c]	4	✓	AA[d], AA, WL[e]
Birney et al, 2016 [31]	Mood Hacker	✓	2	✓	E[f]
Christoforou et al, 2017 [33]	Agoraphobia Free, Stress Free	✓	2	✓	AA
Dahne et al, 2019 [34]	¡Aptívate!, iCouch CBT	✓	3	✓	AA, TAU[g]
Donker et al, 2019 [35]	ZeroPhobia	✓	2	✓	WL
Donovan et al, 2016 [44]	BodiMojo	—	—	—	—
Dulin et al, 2014 [45]	Buddy Steps	—	—	—	—
Hidalgo-Mazzei et al, 2018 [46]	OpenSIMPLe	—	—	—	—
Horsch et al, 2017 [36]	Sleepcare app	✓	2	✓	WL
Hur et al, 2018 [37]	Todac Todac app	✓	2	✓	EMA[h]
Levin et al, 2018 [38]	—	✓	3	✓	AA, WL
Morris et al, 2010 [47]	Mood Map	—	—	—	—
Morrison Wylde et al, 2017 [48]	Headspace	—	2	✓	RLT[i]
Prada et al, 2016 [49]	EMOTEO	—	—	—	—
Pratap et al, 2018 [30]	iPST, EVO, Health Tips	✓	3	✓	AA, AA
Roncero et al, 2018 [50]	GGRO	—	—	—	—
Roy et al, 2015 [39]	GETSmart	✓	2	✓	E
Schlosser et al, 2018 [42]	PRIME-D	—	3	—	AA, AA
Shrier et al, 2017 [52]	—	—	—	—	—
Stoll et al, 2017 [51]	REACH app	—	—	—	—
Versluis et al, 2018 [40]	—	✓	3	✓	WL, EMA
Watts et al, 2013 [41]	Get Happy	✓	2	✓	C[j]
Wen et al, 2017 [53]	Headspace	—	—	—	—

[a]RCT: randomized controlled trial.

[b]Not applicable or no information.

[c]Used in the study design.

[d]AA: another app with different content.

[e]WL: waiting list.

[f]E: encouragement to visit helpful websites or call hotline in case of a difficult situation.

[g]TAU: treatment as usual.

[h]EMA: app with ecological momentary assessment.

[i]RTL: real-life training with a qualified person.

[j]C: computer-based treatment.

Target Populations and Mental Health Problems

EMIs were implemented in nonclinical populations, including healthy participants, and participants with subthreshold symptoms of a mental disorder (10 studies). In these studies, EMIs aimed to tackle a variety of problems, including high self-criticism or high levels of stress in the workplace (eg, of medical doctors, novice nurses, and corporation workers), all of which were assumed to increase the risk of developing a mental health disorder. In additional studies with healthy populations, EMIs addressed subthreshold posttraumatic stress disorder. EMIs were also implemented for the treatment of mental health issues in clinical populations (14 studies), including patients with depression, insomnia, bipolar disorder, schizophrenia, alcohol addiction, borderline personality, agoraphobia, or obsessive-compulsive disorder (Table 2).

Table 2. Sample characteristics.

Study	Population	Final sample size (n)	Power analysis	Age (years)	Age (years), mean (SD)	Gender (female %)
Arevian et al, 2018 [43]	Healthy	28	—[a]	18+	—	40 (11)
Bakker et al, 2018 [32]	Healthy and subthreshold depression and anxiety in screening	198	✓[b]	18-76	34.2 (14.2)	72.7 (144)
Birney et al, 2016 [31]	Clinical: depression	300	✓	18+	—	—
Christoforou et al, 2017 [33]	Clinical: agoraphobia	142	✓	18+	—	—
Dahne et al, 2019 [34]	Clinical: depression	42	—	18+	36.05 (11.44)	67 (28)
Donker et al, 2019 [35]	Clinical: acrophobia	193	✓	18-65	41.33 (13.68)	66.9 (129)
Donovan et al, 2016 [44]	Healthy	16	—	13-22	16.9 (1.3)	65 (10)
Dulin et al, 2014 [45]	Clinical: alcohol use disorder	28	—	22-45	33.6 (5.6)	46 (13)
Hidalgo-Mazzei et al, 2018 [46]	Clinical: bipolar disorder	103	—	18+	36.59 (11)	63.2 (65)
Horsch et al, 2017 [36]	Clinical: insomnia	151	✓	18-80	39.66 (13.44)	62.3 (94)
Hur et al, 2018 [37]	Clinical: depression	34	—	18-35	23.71 (3.26)	88 (30)
Levin et al, 2018 [38]	Healthy: high self-criticism	87	—	18-52	22.76 (7.02)	69 (60)
Morris et al, 2010 [47]	Healthy: high level of stress	8	—	30-48	37 (5.75)	—
Morrison Wylde et al, 2017 [48]	Healthy: medical workers	95	✓	23-50	—	92 (87)
Prada et al, 2016 [49]	Clinical: borderline	16	—	18-50	30.5 (93)	100 (16)
Pratap et al, 2018 [30]	Clinical: depression	345	—	18-70	34.9 (10.72)	77.1 (266)
Roncero et al, 2018 [50]	Clinical: obsessive-compulsive disorder	20	—	19-26	—	80 (16)
Roy et al, 2015 [39]	Subthreshold PTSD[c]	13	—	—	—	—
Schlosser et al, 2018 [42]	Clinical: depression	36	—	18+	31.33 (12.4)	78 (28)
Shrier et al, 2017 [52]	Clinical: depression	16	—	16-23	19.6 (—)	100 (16)
Stoll et al, 2017 [51]	Healthy	132	—	8-12	9.65 (0.82)	62.9 (83)
Versluis et al, 2018 [40]	Healthy: high level of stress	128	✓	18+	—	—
Watts et al, 2013 [41]	Clinical: depression	35	—	18-63	41 (12.38)	80 (28)
Wen et al, 2017 [53]	Healthy: medical workers	30	—	—	—	90 (27)

[a]No information.

[b]Power analysis was conducted.

[c]PTSD: posttraumatic stress disorder.

Sample sizes varied widely, from n=8 to n=348. Seven studies reported a power analysis to justify their sample size. The age range varied from 8 to 80 years. Over one-third of the studies did not report the age range of participants, informing only that they were aged 18 years or older. Five studies did not report the sex composition of their sample. Those that did included a majority (>70%) of female participants. This was due, in part, to the populations of interest. For instance, 1 study was conducted in novice nurses—a profession that is dominated by women. Two other studies included only women. In summary, the sample characteristics of current EMI studies are heterogeneous, and the quality of sample descriptions varies widely across studies.

Content

Cognitive behavioral techniques provide a range of possible interventions that address diverse and complex patient needs [54]. CBT comprises techniques addressing psychological mechanisms that underpin negative and potentially harmful thoughts and beliefs (eg, reflection and cognitive restructuring) and actions (eg, behavioral activation or social skills training) endorsed by patients.

The most frequently used CBT strategies were cognitive restructuring, including reappraisal (14 studies), self-monitoring (13 studies), reflection (10 studies), and relaxation (8 studies). Cognitive restructuring identifies and changes the maladaptive thoughts and beliefs and reevaluates a given situation [15].

Self-monitoring is usually implemented in apps that contain an EMA module and is often described as a key component of successful therapy for many mental health conditions [55,56]. By getting access to their self-reported data, users become self-aware; manage their symptoms; better understand their mood, behavior, or illness; and work on these factors of mental well-being. Reflection is a metacognitive process that allows an individual to increase his or her psychological mindedness and another central process in CBT [57,58]. The reflection process requires the user to look at her or his experiences from a distance and often poses a starting point for many other therapeutic strategies (eg, reappraisal and gratitude). In addition, reflection was used, for instance, as an outcome of self-monitoring [37,44].

The use of digital technologies opens a new door to deliver CBT strategies. Apart from basic solutions, such as sending messages [52] or text-based scenarios [37,50], novel approaches were employed in the reviewed studies. For instance, reflection and cognitive restructuring were presented in a fabular comic story of the main character who had depression [41]. A similar solution was employed in the Agoraphobia free app, which was game based and presented a virtual character who needed to meet the virtual therapist to work on reflection and cognitive restructuring [33]. Relaxation and meditation exercises were delivered via audio and video tools [49,53]. Self-monitoring outcomes were presented, for instance, as *mood cloud*, providing a visual representation of the participant's self-reported mood [44], or in a calendar view, to allow the user to track their behavior day-by-day or even hour-by-hour [36]. One novel solution combines a mobile app with virtual reality (VR) to treat acrophobia with exposure [35] or a platform allowing users to contact each other to provide social support that enhances behavioral activation [42]. All apps in the studies reviewed here used more than one strategy to improve mental health (Table 3).

Table 3. Cognitive behavioral therapy techniques implemented in ecological momentary interventions.

Study	BA[a]	CR[b]	DF[c]	DSC[d]	DST[e]	EXP[f]	MDF[g]	PS[h]	PE[i]	RF[j]	RLX[k]	SMN[l]	SST[m]
Arevian et al, 2018 [43]	—[n]	✓[o]	—	—	—	—	—	—	—	—	—	—	✓
Bakker et al, 2018 [32][p]	✓	—	—	—	—	—	—	—	✓	—	—	✓	—
Birney et al, 2016 [31]	✓	✓	—	—	—	—	✓	—	—	—	—	✓	—
Christoforou et al, 2017 [33]	—	✓	—	—	✓	✓	—	—	✓	✓	✓	—	—
Dahne et al, 2019 [34]	✓	—	—	—	—	—	—	—	✓	✓	—	✓	—
Donker et al, 2019 [35][p]	—	—	—	—	—	✓	—	—	✓	—	—	—	—
Donovan et al, 2016 [44]	—	✓	—	—	—	—	✓	—	—	✓	—	✓	—
Dulin et al, 2014 [45]	✓	—	—	—	—	—	—	✓	✓	—	—	✓	✓
Hidalgo-Mazzei et al, 2018 [46]	—	—	—	—	—	—	—	—	✓	—	—	✓	—
Horsch et al, 2017 [36]	✓	—	—	—	—	—	—	—	✓	—	✓	✓	—
Hur et al, 2018 [37]	—	✓	—	✓	—	—	—	—	—	✓	—	✓	—
Levin et al, 2018 [38]	—	✓	✓	—	—	—	—	—	—	✓	—	—	—
Morris et al, 2010 [47]	—	✓	—	—	—	—	—	—	—	—	✓	✓	—
Morrison Wylde et al, 2017 [48]	—	—	—	—	—	—	✓	—	—	✓	✓	—	—
Prada et al, 2016 [49]	—	—	—	—	✓	—	✓	—	—	—	✓	✓	—
Pratap et al, 2018 [30]	—	✓	—	—	—	—	—	✓	✓	—	—	—	—
Roncero et al, 2018 [50]	✓	✓	—	—	—	—	—	—	—	—	—	—	—
Roy et al, 2015 [39]	—	✓	—	—	—	—	—	—	✓	✓	✓	—	✓
Schlosser et al, 2018 [42]	✓	—	—	—	—	—	✓	—	✓	—	—	—	✓
Shrier et al, 2017 [52]	—	✓	—	—	—	—	—	—	—	✓	—	✓	—
Stoll et al, 2017 [51]	—	✓	—	—	—	—	—	—	—	—	✓	✓	—
Versluis et al, 2018 [40]	—	✓	—	—	—	—	✓	—	—	—	—	✓	—
Watts et al, 2013 [41]	—	✓	—	—	—	—	—	—	✓	✓	—	—	✓
Wen et al, 2017 [53]	—	—	—	—	—	—	✓	—	—	✓	✓	—	—

[a]BA: behavioral activation.

[b]CR: cognitive restructuring, including reappraisal.

[c]DF: defusion.

[d]DSC: distancing.

[e]DST: distraction.

[f]EXP: exposure.

[g]MDF: mindfulness.

[h]PS: problem solving.

[i]PE: psychoeducation.

[j]RF: reflection.

[k]RLX: relaxation.

[l]SMN: self-monitoring.

[m]SST: social skills training.

[n]Technique not implemented in the ecological momentary intervention.

[o]Technique implemented in the ecological momentary intervention.

[p]Insufficient information was provided—probably more modules were implemented.

Delivery

The most common way to deliver EMIs on a smartphone is a mobile app, but ever-growing possibilities are emerging with technological advances. For example, mobile apps can be combined with wristbands to track indicators of physiological function (eg, heartbeat), GPS to track geolocation, or VR.

The duration of interventions ranged from 2 weeks to 6 months, with a median of 30 days. A frequent option was to send prompts

to remind participants either about using EMI or about completing the EMA (17 studies). The number of prompts with EMA or reminders varied from 5 per day to once every week (Table 4). In 4 apps, prompts were sent only when the participant forgot to use the app for a longer, predefined period. Another way to deliver EMIs was to allow participants to use them anytime they needed to or once per day.

Table 4. Characteristics of ecological momentary intervention delivery.

Study	Intervention duration in days	Prompts sent	Number of prompts per day	Trigger
Arevian et al, 2018 [43]	28	✓[a]	1	Time based
Bakker et al, 2018 [32]	30	✓ 1/3[b]	N/I[c]	N/A[d]
Birney et al, 2016 [31]	42	✓	1	N/I
Christoforou et al, 2017 [33]	84	✓	0.14	N/I
Dahne et al, 2019 [34]	56	✓	0.14	Participant
Donker et al, 2019 [35]	21	✓	0.14	Participant
Donovan et al, 2016 [44]	30	N/A	N/A	Participant
Dulin et al, 2014 [45]	42	✓	1	Random
Hidalgo-Mazzei et al, 2018 [46]	168	✓	1	EMA[e]
Horsch et al, 2017 [36]	84	✓	1 AL[f]	Depending on data[g]
Hur et al, 2018 [37]	21	✓	0.5 ANU[h]	EMA
Levin et al, 2018 [38]	14	✓	4	Participant and random
Morris et al, 2010 [47]	30	✓	N/I	Participant
Morrison Wylde et al, 2017 [48]	28	N/A	N/A	N/A
Prada et al, 2016 [49]	168	N/A	N/A	N/A
Pratap et al, 2018 [30]	84	✓	1	Participant
Roncero et al, 2018 [50]	15	✓	1	N/I
Roy et al, 2015 [39]	42	✓	1	Depending on data
Schlosser et al, 2018 [42]	56	N/A	N/A	Participant
Shrier et al, 2017 [52]	28	✓	5	Random
Stoll et al, 2017 [51]	1	N/A	N/A	N/A
Versluis et al, 2018 [40]	26	✓	5	Random
Watts et al, 2013 [41]	56	N/A	N/A	N/A
Wen et al, 2017 [53]	30	N/A	N/A	Participant

[a]Prompts were sent by ecological momentary intervention.

[b]1/3: in 1 of the 3 apps introduced in the study (when not marked, all apps introduced in the study were sending beeps).

[c]N/I: no information.

[d]N/A: not applicable.

[e]EMA: ecological momentary assessment.

[f]AL: at least (depending on participants' reports).

[g]Depending on data: depending on participants' report other than ecological momentary assessment (eg, sleep hygiene).

[h]ANU: beeps were sent only when app was not used.

Triggers can be divided into 3 categories: time based, event based, and randomized within a specific time frame. In 10 studies, event-based triggers were employed to allow participants to trigger EMI themselves. For example, participants were instructed to use the app anytime they felt they might benefit from the EMI (Table 4). Another trigger from this category was EMA (used in 2 studies): when problems or symptoms were reported (eg, low mood or above-threshold stress level), the participant received an EMI. This just-in-time adaptive intervention (JITAI) aims to provide the right type or amount of support, at the right time, by adapting to an individual's changing internal and contextual state [59]. Four studies used randomized prompt-sending as an EMI trigger to avoid prepared answers from participants. One app sent prompts every morning.

The EMI frequency varied among studies. The range of the frequency of EMI delivery was from 3 times per day to once

per week. Four studies did not report the number of delivered EMIs.

Taken together, the frequency, triggers, and duration of the intervention varied and depended on the target group, targeted problem, and CBT techniques implemented. A few studies did not report this information, including default or the average number of prompts per participant or duration of intervention for different participants.

Feasibility

For this review, feasibility was indexed by compliance rate and participants' satisfaction with the intervention.

Compliance reporting was variable. First, the definition of compliance differed across studies. Some researchers reported how many training sessions participants completed, how many prompts participants answered, or the number of participants who did not drop out of the study. Others set thresholds, such as reacting to at least one or two prompts and subsequently excluded all participants with lower adherence. For the purpose of this review, we defined *compliance rate* as the number of reactions to prompts. Nineteen studies did not report compliance. None of the studies with healthy samples reported compliance rates.

In other studies, compliance ranged from 33.8% to 93.3% (7 studies; mean 64, SD 22). There was no relationship between compliance rate and duration as well as the overall time cost of the intervention. For instance, one study reported a compliance rate of 55.6% for a 15-day intervention with 1 prompt per day [47], and another reported a compliance rate of 33.8% for a 168-day intervention with 1 prompt per day [43], whereas other studies reported a compliance rate of 90.5% for a 56-day intervention with 1 prompt per week [31] or a compliance rate of 93.3% for a 42-day intervention with 1 prompt per day [28]. All studies with higher than average compliance rates included participants with depression. Lower compliance rates were reported in a study of participants with bipolar disorder.

Satisfaction and helpfulness rates were underreported (satisfaction was reported in 7 studies and helpfulness in 4 studies) but rated positively (mean perceived satisfaction 72.6%, SD 17.2 and mean perceived helpfulness 70.8%, SD 15.3). Helpfulness was rated highest by individuals with borderline personality disorder and healthy samples with high self-criticism. Healthy populations, especially adolescents, were generally more satisfied with the EMIs compared with clinical populations (Table 5). The least satisfied were users with bipolar disorder; however, only 1 study was conducted on such a sample. Discrepancies were found in satisfaction rates reported by depressed samples, ranging from 46% and 54% (which were the lowest reported rates) to 91.8% (which was one of the highest rates). However, only 3 of 7 studies that delivered EMI to depressed populations reported these numbers.

Table 5. Feasibility, efficacy, and effectiveness.

Study	Compliance in %	Helpfulness in %	Satisfaction in %	Primary outcome measure	Improvement in primary outcome	Secondary outcome measure
Arevian et al, 2018 [43]	__[a]	—	—	—	—	—
Bakker et al, 2018 [32]	—	—	—	PHQ[b], GAD[c], WEMWBS[d]	✓[e]	ESAS-R[f], CSES[g], MHLQ[h]
Birney et al, 2016 [31]	93	—	46	PHQ	—	BADS[i], ATQ-R[j], KT[k]
Christoforou et al, 2017 [33]	52	—	—	PAS[l]	✓	F[m]
Dahne et al, 2019 [34]	91	—	—	PHQ	✓	F
Donker et al, 2019 [35]	—	—	—	AQ[n]	✓	ATHQ[o], BDI-II[p], IPQ[q], PHQ, M[r], SUS[s]
Donovan et al, 2016 [44]	—	64	92	—	—	—
Dulin et al, 2014 [45]	—	—	—	TLFB[t]	✓	F
Hidalgo-Mazzei et al, 2018 [46][u]	33.8	—	62	WHO-5[v], SF-36[w]	✓	F
Horsch et al, 2017 [36]	49.7	—	—	ISI[x]	✓	PSQI[y], DBAS-16[z], HADS[aa], CES-D[ab]
Hur et al, 2018 [37]	—	—	—	DAS[ac], BDI-II, STAI[ad], RSES[ae]	✓	QoL[af]
Levin et al, 2018 [38]	—	78 and 83	81 and 81	FSCRS[ag]	✓	DASS[ah], WSAS[ai]
Morris et al, 2010 [47]	—	—	—	I[aj], EMASS[ak]	✓	—
Morrison Wylde et al, 2017 [48]	—	—	—	CFST[al]	✓	LEC[am], PCL-C[an], FFMQ[ao]
Prada et al, 2016 [49]	—	82	—	EMAAT[ap]	✓	—
Pratap et al, 2018 [30][u]	—	—	—	PHQ, SDS[aq]	—	CR[ar]
Roncero et al, 2018 [50]	56	—	—	OCI-R[as], OBQ[at]	✓	ROCI[au], PROC-SI[av], DASS
Roy et al, 2015 [39]	—	—	—	PCL[aw]	✓	PHQ, GAD
Schlosser et al, 2018 [42]	—	—	72	PHQ, SDS	✓	—
Shrier et al, 2017 [52]	—	—	—	—	—	—
Stoll et al, 2017 [51]	—	—	92	—	—	—
Versluis et al, 2018 [40]	—	47	—	HR[ax]	—	IPANAT[ay], IAT[az], ERI[ba], PSWQ[bb], GAD, PHQ, FFMQ, CEQ[bc], EMASE[bd], F
Watts et al, 2013 [41]	69	—	54	PHQ	✓	K-10[be], BDI-II, CEQ, SDS, ERS[bf], HRS[bg]

Study	Compliance in %	Helpfulness in %	Satisfaction in %	Primary outcome measure	Improvement in primary outcome	Secondary outcome measure
Wen et al, 2017 [53]	—	—	—	PANAS[bh], FMI[bi]	✓	—

[a]No information.

[b]PHQ: Patient Health Questionnaire.

[c]GAD: Generalized Anxiety Disorder Scale.

[d]WEMWBS: Warwick-Edinburgh Mental Well-Being Scale.

[e]Study with improvement in the primary outcomes. Improvement in primary outcomes=results statistically significant in at least one measure of the primary outcomes.

[f]ESAS-R: Emotional Self-Awareness Scale-Revised.

[g]CSES: Coping Self-Efficacy Scale.

[h]MHLQ: Mental Health Literacy Questionnaire.

[i]BADS: Behavioral Activation for Depression Scale.

[j]ATQ-R: Automatic Thoughts Questionnaire-Revised.

[k]KT: knowledge test.

[l]PAS: Panic and Agoraphobia Scale.

[m]F: feasibility measures.

[n]AQ: Acrophobia Questionnaire.

[o]ATHQ: Attitudes Toward Heights Questionnaire.

[p]BDI-II: Beck Depression Inventory II.

[q]IPQ: Igroup Presence Questionnaire.

[r]M: mastery.

[s]SUS: System Usability Scale.

[t]TLFB: The Timeline Followback.

[u]To provide information about efficacy, we reversed the order of the outcomes.

[v]WHO-5: World Health Organization 5-point Well-Being Index.

[w]SF-36: Short Form Health Survey.

[x]ISI: Insomnia Severity Index.

[y]PSQI: Pittsburgh Sleep Quality Index.

[z]DBAS-16: Dysfunctional Beliefs and Attitudes about Sleep.

[aa]HADS: Hospital Anxiety and Depression Scale.

[ab]CES-D: Center of Epidemiologic Studies Depression Scale.

[ac]DAS: Dysfunctional Attitude Scale.

[ad]STAI: State-Trait Anxiety Inventory.

[ae]RSES: Rosenberg Self-Esteem Scale.

[af]QoL: quality of life.

[ag]FSCRS: Forms of Self-Criticism and Self-Reassurance Scale.

[ah]DASS: Depression, Anxiety, and Stress Scale.

[ai]WSAS: Work and Social Adjustment Scale.

[aj]I: interview.

[ak]EMASS: Ecological Momentary Assessment-Stress Scale.

[al]CFST: Compassion Fatigue Self-Test.

[am]LEC: Life Events Checklist.

[an]PCL-C: Posttraumatic Stress Disorder Checklist-Civilian Version.

[ao]FFMQ: Five Facet Mindfulness Questionnaire.

[ap]EMAAT: Ecological Momentary Assessment-Aversive Tension Scale.

[aq]SDS: Sheenan Disability Scale.

[ar]CR: comparison of recruitment and engagement in a fully remote trial of individuals with depression who either self-identify as Hispanic/Latino or not.

[as]OCI-R: Obsessive-Compulsive Inventory-Revised.

[at]OBQ: Obsessive Beliefs Questionnaire.

[au]ROCI: Relationship Obsessive-Compulsive Inventory.

[av]PROCSI: Partner-Related Obsessive-Compulsive Symptoms Inventory.

[aw]PCL: Posttraumatic Stress Disorder Checklist.
[ax]HR: heart rate.
[ay]IPANAT: Implicit Positive and Negative Affect Test.
[az]IAT: Implicit Association Test.
[ba]ERI: Effort-Reward Imbalance Questionnaire.
[bb]PSWQ: Penn State Worry Questionnaire.
[bc]CEQ: Credibility/Expectancy Questionnaire.
[bd]EMASE: Ecological Momentary Assessment-Stress and Explicit Affect.
[be]K-10: Kessler 10-Item Psychological Distress Scale.
[bf]ERS: Environment Rating Scale.
[bg]HRS: Homework Rating Scale.
[bh]PANAS: Positive and Negative Affect Scale.
[bi]FMI: Freiburg Mindfulness Scale.

Efficacy and Effectiveness

Efficacy or effectiveness was reported in 21 studies. In the 8 studies reviewed here, the authors used both these terms alternately to report the outcomes of EMI. Hence, both effectiveness and efficacy are reported in this review, and the particular term is used as indicated by the authors of the studies.

In total, 16 of the 21 interventions included in our review reported evidence of a significant reduction in mental disorder symptoms (Table 5).

Nine studies reported effect sizes (Cohen *d*). In the EMI study for insomnia, the authors found significant interaction effects between time and condition on the primary outcome measures (*d*=−0.66) and sleep efficiency (*d*=0.71) [33]. A significant effect on insomnia severity was also recorded at the 3-month follow-up. In the study investigating high self-criticism, effect sizes for between-group comparison and for all primary outcome measures ranged from 0.46 to 0.72 (except inadequate self-criticism, which was nonsignificant) [35]. Another study assessing the effectiveness of EMI in individuals with acrophobia found a between-group effect size *d*=1.14 [32].

The within-group effect sizes varied from *d*=0.47 to *d*=0.64 and *d*=0.67 for depressed samples and changes in overall health, as indexed by the Patient Health Questionnaire. One study found effect sizes of *d*=0.77 for measures of mindfulness in a medical workers' sample and a moderate effect size *d*=0.38 for positive affect [50]. A large effect (*d*=1.0) was found in the study on a population with alcohol use disorder [42]. The only effect sizes that were small or null were found in the study investigating EMI helping in obsessive-compulsive disorder [47]. Taken together, the effect sizes for a reduction in primary outcomes mostly varied from moderate to large, with large effects found for the majority of the reviewed EMIs.

The most common method to assess efficacy and effectiveness was with questionnaires, and almost all the studies using this method found statistically significant differences (in both between- and within-group comparisons) in the primary outcomes (Table 5). With respect to other measures of primary outcomes, 1 study used heart rate [37] and 2 other studies calculated effectiveness based on EMA data [44,46]. Although the outcomes based on EMA were statistically significant, there were no effects found based on physiological readouts.

As for secondary outcomes, 1 of the studies employed a knowledge test about depression and found statistically significant positive effects for the program on information and knowledge [28]. In another study, the authors added an Implicit Positive and Negative Affect Test to assess the level of unconscious stress and found that implicit stress did decrease over time for all participants, independent of the intervention [37].

Discussion

Principal Findings

We reviewed 26 studies that investigated the feasibility, efficacy, and effectiveness of standalone CBT-based EMIs to increase mental health. Results show that EMIs can be successfully delivered via smartphones, significantly increase well-being among users, and reduce symptoms of mental disorders. Designs and quality of the studies reviewed using a rational app design were heterogeneous. Across studies, EMIs were generally accepted by users with various age, sex, education background, and professions and were shown to be effective treatments for a broad range of psychological symptoms.

CBT assumptions constitute an evidence-based basis for EMIs, especially with the use of techniques involving self-monitoring, reflection, cognitive restructuring, and relaxation. Such techniques were mostly employed in combination, and users were sometimes provided with a personal choice of which approach they preferred.

Many studies investigated the efficacy of reducing symptoms of depression in clinical populations. The Todac Todac app, for instance, significantly reduced both depression and anxiety symptoms [37]. The Get Happy program was associated with stable reduction of depressive symptoms, including a demonstration of stability at 3 months follow-up [41]. Promising results were also shown in studies with healthy samples, where an increase in positive well-being and higher emotional self-awareness along with self-efficacy as a result of using EMI were reported. Here, the MoodKit, MoodMission, and MoodPrism apps were effective in decreasing mild depressive and anxiety symptoms [32]. Headspace is a mindfulness app that was used in 2 studies with medical workers and tested against a control group that received face-to-face mindfulness classes. Users engaging in Headspace reported higher

self-awareness in their actions, less burnout, and more compassion satisfaction compared with the group receiving face-to-face mindfulness training [47]. In another study of the Headspace app, a single group design was chosen. This study reported heightened mindfulness skills and more positive emotions as a result of using Headspace post-EMI versus pre-EMI in medical workers [53].

Clinical Perspective

From a practical clinical perspective, EMIs may postpone the necessity of professional mental health treatment by a psychiatrist or psychologist, eg, for individuals who struggle financially, have limited access to such services in their place of living, or due to health reasons cannot attend psychotherapy [32,51]. On the basis of this review and previous studies, EMIs may be specifically helpful for individuals who are facing temporary problems, for instance, in the context of personal or professional change or other transitioning periods. Moreover, symptoms of mental health are still affected by considerable stigmatization [60], and EMIs may help reduce the fear of stigma as psychological support can be obtained without seeking treatment in a hospital setting. EMIs are also often a first step for many individuals to test personal preferences and affinity to psychological treatment and may lead to face-to-face or other professional therapeutic services later [61]. The worldwide long-lasting crisis of mental health care, manifested by poor accessibility of cost-free, state-funded psychotherapy, suggests another use of EMIs. Patients could use EMIs as a source of support while awaiting face-to-face treatment [62-64].

EMIs in Context of Pandemics

In the context of COVID-19, the need to implement population-based behavioral measures to counter the rapid spread of the virus has led to home confinement; loss of individual behavioral routines, including exercise; and lack of social contact. Psychological needs are often underaddressed in such times of crises [8,65]. EMIs could be an ideally suited approach to reach a large number of individuals affected by mental health problems and to provide scalable global mental health solutions. It is also estimated, based on experiences with previous pandemics, that after the COVID-19 pandemic is over, such solutions will be needed to help populations with continued psychological symptoms, such as depression and anxiety [66]. Some populations might be at a specific risk of stress in this challenging time, especially doctors, nurses, or emergency personnel [67-69], and 2 of the trials reviewed here investigated effects on the medical staff. The high efficacy of the Headspace app tested in these populations of first responders and frontline workers is encouraging and indicates that mindfulness training delivered via a mobile app can significantly decrease work-related stress. EMIs could thus help to lower the risk of developing unfavorable mental health outcomes in such populations, which have recently been documented [70].

Methodological Limitations

There are several methodological caveats in the studies reviewed here. The choice of control condition varied. Suggestions for the control groups to call hotlines or visit recommended websites imply heterogeneous and less tangible control groups than control groups that could use comparison apps. At the same time, comparing results from 2 different apps can be difficult if they target different symptoms or use different techniques. Little attention has been paid to the psychological mechanisms (eg, cognitive mediators) that allowed for this change in the outcome to develop. To increase the conclusiveness and reliability of studies assessing efficacy, effectiveness, and feasibility of EMIs, more RCTs with adequate control groups are needed [71]. Ideally, RCTs should be preceded by optimization studies, based, for instance, on MOST (Multiphase Optimization Strategy) [72] and SMART (The Sequential Multiple Assignment Randomized Trial) [73], which allow for choosing the right dosage and delivery of intervention based on predefined criteria, for example, individual characteristics of the user, characteristics of psychiatric population, and effectiveness of the intervention. Another valid concern is the reporting of EMI studies. For example, the majority (73%) of the studies did not report a formal power analysis. Consequently, studies might be underpowered, and their quantitative results have not been reported. However, some studies investigated the feasibility rather than efficacy and focused on issues of implementation, launch, and users' experience with the app; thus, they might have been less concerned about power and sample size. Moreover, it must be noted that there might also be a bias in overestimating the effects of EMI. First, control conditions, if used, cannot simply be blinded, similar to psychotherapy studies. Second, the fact that almost all studies showed a positive effect on the primary outcome makes it likely that there is a publication bias toward positive results.

Despite these limitations, we indicated key strengths and found a number of standalone smartphone-based EMIs that were shown to be effective. It is also in line with findings from previous reviews in this field, although their focus differed, for instance, by focusing on blended therapies [24,25] or specific mental health problems [12,19-23].

Recommendations for Future Development

Exciting future developments lie ahead in the field of EMIs. This is documented by >40 (pre)registered protocols, which we excluded (Figure 1) and which are likely to produce further evidence about the effectiveness of CBT-based standalone EMIs. Future developments should exploit the potential to scale these interventions and test their effects in larger and global populations in times of crises, such as the COVID-19 pandemic. Moreover, it is possible to combine EMIs with wearable technologies [74] to gather physiological data and other information for researchers and clinicians alike. EMIs can be used along passive sensing tracking patterns of smartphone usage, and by using context-aware assessment strategies that link the assessment of experiences to specific sensing events (like, for instance, physical distance to the others), would allow for the design of specific interventions in response to COVID-19 pandemic–induced mental health problems, especially induced by isolation whereas, at the same time, complying with the demanded pandemic containment policies. Here, mobile crowdsensing in combination with EMI would constitute the key technology to address this major mental health challenge. Ideally, these tools can be combined with machine learning to derive algorithms for accurate and precise prediction and

selection of ecologically valid treatment options [75]. Machine learning, or more generally artificial intelligence techniques, could be useful in at least three ways. First, such approaches could identify individuals and/or time points where the intervention might be particularly important to avoid substantial degradation of mental health.

Second, such approaches could identify factors that critically affect effectiveness, which could subsequently help in personalizing EMIs both delivery mode and contentwise, and in enhancing decision making at precise time points while using personally tailored JITAIs.

Third, they could significantly decrease the time invested by health professionals and scientists who administer EMIs, such as in the study where clinicians manually adapted the number of EMIs delivered based on app usage [42].

Conclusions

Taken together, our review showed that standalone CBT-based EMIs can be efficacious in reducing symptoms of psychological disorders in various situations. EMIs may support individuals outside the therapists' room and could also be helpful when time and resource investment from mental health professionals are limited. At the same time, however, professional clinician-scientists' input is invaluable in designing and implementing such EMIs. All EMIs reviewed here were based on a rational app design that translates key concepts and findings from CBT to these new interventions. They can provide scalable and evidence-based mental health support for large populations and be readily distributed.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 777084. This study reflects only the authors' view and the European Commission is not responsible for any use that may be made of the information it contains.

Authors' Contributions

MM, LS, and BK designed the study. Screening, full text review, and data extraction were conducted by MM (100%) and checked (100%) by JR. Quality assessment of the studies was conducted by MM and CW. MM wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Complete search strings.
[DOC File , 25 KB - mhealth_v8i11e19836_app1.doc]

References

1. Mental Disorders Affect One in Four People. World Health Organization. URL: https://www.who.int/whr/2001/media_centre/press_release/en/ [accessed 2020-10-01]
2. Copeland W, Shanahan L, Costello EJ, Angold A. Cumulative prevalence of psychiatric disorders by young adulthood: a prospective cohort analysis from the great smoky mountains study. J Am Acad Child Adolesc Psychiatry 2011 Mar;50(3):252-261 [FREE Full text] [doi: 10.1016/j.jaac.2010.12.014] [Medline: 21334565]
3. Moffitt TE, Caspi A, Taylor A, Kokaua J, Milne BJ, Polanczyk G, et al. How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. Psychol Med 2010 Jun;40(6):899-909 [FREE Full text] [doi: 10.1017/S0033291709991036] [Medline: 19719899]
4. Ritchie H, Roser M. Mental Health. Our World Data Internet. 2018. URL: https://ourworldindata.org/mental-health [accessed 2020-10-01]
5. Pfefferbaum B, Schonfeld D, Flynn B, Norwood AE, Dodgen D, Kaul RE, et al. The H1N1 crisis: a case study of the integration of mental and behavioral health in public health crises. Disaster Med Public Health Prep 2012 Mar;6(1):67-71. [doi: 10.1001/dmp.2012.2] [Medline: 22490939]
6. Reissman D, Watson P, Klomp R, Tanielian T, Prior S. Pandemic influenza preparedness: adaptive responses to an evolving challenge. J Homel Secur Emerg Manag 2006;3(2):- [FREE Full text] [doi: 10.2202/1547-7355.1233]
7. Brooks S, Webster R, Smith L, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar;395(10227):912-920 [FREE Full text] [doi: 10.1016/s0140-6736(20)30460-8]
8. Duan L, Zhu G. Psychological interventions for people affected by the COVID-19 epidemic. Lancet Psychiatry 2020 Apr;7(4):300-302 [FREE Full text] [doi: 10.1016/S2215-0366(20)30073-0] [Medline: 32085840]

9. Mental Disorders. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/mental-disorders [accessed 2020-10-01]
10. Wang PS, Angermeyer M, Borges G, Bruffaerts R, Tat Chiu W, de Girolamo G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007 Oct;6(3):177-185. [Medline: 18188443]
11. Donker T, Cornelisz I, van Klaveren C, van Straten A, Carlbring P, Cuijpers P, et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. JAMA Psychiatry 2019 Jul 1;76(7):682-690 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.0219] [Medline: 30892564]
12. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Mar;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
13. Wilhelm FH, Grossman P. Emotions beyond the laboratory: theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment. Biol Psychol 2010 Jul;84(3):552-569. [doi: 10.1016/j.biopsycho.2010.01.017] [Medline: 20132861]
14. Myin-Germeys I, Klippel A, Steinhart H, Reininghaus U. Ecological momentary interventions in psychiatry. Curr Opin Psychiatry 2016 Jul;29(4):258-263. [doi: 10.1097/YCO.0000000000000255] [Medline: 27153125]
15. Beck AT. Cognitive therapy: nature and relation to behavior therapy. Behav Ther 1970 May;1(2):184-200. [doi: 10.1016/s0005-7894(70)80030-2]
16. Ellis A. Reason and Emotion in Psychotherapy. Oxford, England: Lyle Stuart; 1962.
17. Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. Clin Psychol Rev 2006 Jan;26(1):17-31. [doi: 10.1016/j.cpr.2005.07.003] [Medline: 16199119]
18. David D, Cristea I, Hofmann SG. Why cognitive behavioral therapy is the current gold standard of psychotherapy. Front Psychiatry 2018;9:4 [FREE Full text] [doi: 10.3389/fpsyt.2018.00004] [Medline: 29434552]
19. Loo Gee B, Griffiths KM, Gulliver A. Effectiveness of mobile technologies delivering ecological momentary interventions for stress and anxiety: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):221-229. [doi: 10.1093/jamia/ocv043] [Medline: 25997643]
20. Versluis A, Verkuil B, Spinhoven P, van der Ploeg MM, Brosschot JF. Changing mental health and positive psychological well-being using ecological momentary interventions: a systematic review and meta-analysis. J Med Internet Res 2016 Jun 27;18(6):e152 [FREE Full text] [doi: 10.2196/jmir.5642] [Medline: 27349305]
21. Schueller SM, Aguilera A, Mohr DC. Ecological momentary interventions for depression and anxiety. Depress Anxiety 2017 Jun;34(6):540-545. [doi: 10.1002/da.22649] [Medline: 28494123]
22. Anastasiadou D, Folkvord F, Lupiañez-Villanueva F. A systematic review of mHealth interventions for the support of eating disorders. Eur Eat Disord Rev 2018 Sep;26(5):394-416. [doi: 10.1002/erv.2609] [Medline: 29927004]
23. Lan A, Lee A, Munroe K, McRae C, Kaleis L, Keshavjee K, et al. Review of cognitive behavioural therapy mobile apps using a reference architecture embedded in the patient-provider relationship. Biomed Eng Online 2018 Dec 17;17(1):183 [FREE Full text] [doi: 10.1186/s12938-018-0611-4] [Medline: 30558610]
24. Erbe D, Eichert H, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Sep 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
25. Rathbone AL, Clarry L, Prescott J. Assessing the efficacy of mobile health apps using the basic principles of cognitive behavioral therapy: systematic review. J Med Internet Res 2017 Nov 28;19(11):e399 [FREE Full text] [doi: 10.2196/jmir.8598] [Medline: 29187342]
26. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions. New York, USA: The Cochrane Collaboration; 2011.
27. Ryan R, Hill S, Prictor M. Cochrane Consumers and Communication Review Group. Cochrane Consumers and Communication. URL: http://cccrg.cochrane.org/authorresources [accessed 2020-09-30]
28. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]
29. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]
30. Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using mobile apps to assess and treat depression in Hispanic and Latino populations: fully remote randomized clinical trial. J Med Internet Res 2018 Aug 9;20(8):e10130 [FREE Full text] [doi: 10.2196/10130] [Medline: 30093372]
31. Birney AJ, Gunn R, Russell JK, Ary DV. MoodHacker mobile web app with email for adults to self-manage mild-to-moderate depression: randomized controlled trial. JMIR Mhealth Uhealth 2016 Jan 26;4(1):e8 [FREE Full text] [doi: 10.2196/mhealth.4231] [Medline: 26813737]

32. Bakker D, Kazantzis N, Rickwood D, Rickard N. A randomized controlled trial of three smartphone apps for enhancing public mental health. Behav Res Ther 2018 Oct;109:75-83. [doi: 10.1016/j.brat.2018.08.003] [Medline: 30125790]
33. Christoforou M, Sáez Fonseca JA, Tsakanikos E. Two novel cognitive behavioral therapy-based mobile apps for agoraphobia: randomized controlled trial. J Med Internet Res 2017 Nov 24;19(11):e398 [FREE Full text] [doi: 10.2196/jmir.7747] [Medline: 29175809]
34. Dahne J, Collado A, Lejuez CW, Risco CM, Diaz VA, Coles L, et al. Pilot randomized controlled trial of a Spanish-language behavioral activation mobile app (¡Aptívate!) for the treatment of depressive symptoms among united states Latinx adults with limited English proficiency. J Affect Disord 2019 May 1;250:210-217 [FREE Full text] [doi: 10.1016/j.jad.2019.03.009] [Medline: 30870770]
35. Donker T, Cornelisz I, van Klaveren C, van Straten A, Carlbring P, Cuijpers P, et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. JAMA Psychiatry 2019 Jul 1;76(7):682-690 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.0219] [Medline: 30892564]
36. Horsch CH, Lancee J, Griffioen-Both F, Spruit S, Fitrianie S, Neerincx MA, et al. Mobile phone-delivered cognitive behavioral therapy for insomnia: a randomized waitlist controlled trial. J Med Internet Res 2017 Apr 11;19(4):e70 [FREE Full text] [doi: 10.2196/jmir.6524] [Medline: 28400355]
37. Hur J, Kim B, Park D, Choi S. A scenario-based cognitive behavioral therapy mobile app to reduce dysfunctional beliefs in individuals with depression: a randomized controlled trial. Telemed J E Health 2018 Sep;24(9):710-716. [doi: 10.1089/tmj.2017.0214] [Medline: 29323626]
38. Levin ME, Haeger J, An W, Twohig MP. Comparing cognitive defusion and cognitive restructuring delivered through a mobile app for individuals high in self-criticism. Cogn Ther Res 2018 Jul 13;42(6):844-855. [doi: 10.1007/s10608-018-9944-3]
39. Roy MJ, Costanzo ME, Highland KB, Olsen C, Clayborne D, Law W. An app a day keeps the doctor away: guided education and training via smartphones in subthreshold post traumatic stress disorder. Cyberpsychol Behav Soc Netw 2017 Aug;20(8):470-478. [doi: 10.1089/cyber.2017.0221] [Medline: 28737954]
40. Versluis A, Verkuil B, Spinhoven P, Brosschot J. Effectiveness of a smartphone-based worry-reduction training for stress reduction: a randomized-controlled trial. Psychol Health 2018 Sep;33(9):1079-1099. [doi: 10.1080/08870446.2018.1456660] [Medline: 29609475]
41. Watts S, Mackenzie A, Thomas C, Griskaitis A, Mewton L, Williams A, et al. CBT for depression: a pilot RCT comparing mobile phone vs. computer. BMC Psychiatry 2013 Mar 7;13:49 [FREE Full text] [doi: 10.1186/1471-244X-13-49] [Medline: 23391304]
42. Schlosser DA, Campellone TR, Truong B, Anguera JA, Vergani S, Vinogradov S, et al. The feasibility, acceptability, and outcomes of PRIME-D: a novel mobile intervention treatment for depression. Depress Anxiety 2017 Jun;34(6):546-554 [FREE Full text] [doi: 10.1002/da.22624] [Medline: 28419621]
43. Arevian AC, O'Hora J, Jones F, Mango J, Jones L, Williams PG, et al. Participatory technology development to enhance community resilience. Ethn Dis 2018;28(Suppl 2):493-502 [FREE Full text] [doi: 10.18865/ed.28.S2.493] [Medline: 30202203]
44. Donovan E, Rodgers RF, Cousineau TM, McGowan KM, Luk S, Yates K, et al. Brief report: feasibility of a mindfulness and self-compassion based mobile intervention for adolescents. J Adolesc 2016 Dec;53:217-221. [doi: 10.1016/j.adolescence.2016.09.009] [Medline: 27814499]
45. Dulin PL, Gonzalez VM, Campbell K. Results of a pilot test of a self-administered smartphone-based treatment system for alcohol use disorders: usability and early outcomes. Subst Abus 2014;35(2):168-175 [FREE Full text] [doi: 10.1080/08897077.2013.821437] [Medline: 24821354]
46. Hidalgo-Mazzei D, Reinares M, Mateu A, Nikolova VL, Bonnín CD, Samalin L, et al. OpenSIMPLe: a real-world implementation feasibility study of a smartphone-based psychoeducation programme for bipolar disorder. J Affect Disord 2018 Dec 1;241:436-445. [doi: 10.1016/j.jad.2018.08.048] [Medline: 30145515]
47. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. J Med Internet Res 2010 Apr 30;12(2):e10 [FREE Full text] [doi: 10.2196/jmir.1371] [Medline: 20439251]
48. Wylde CM, Mahrer NE, Meyer RM, Gold JI. Mindfulness for novice pediatric nurses: smartphone application versus traditional intervention. J Pediatr Nurs 2017;36:205-212. [doi: 10.1016/j.pedn.2017.06.008] [Medline: 28888505]
49. Prada P, Zamberg I, Bouillault G, Jimenez N, Zimmermann J, Hasler R, et al. Emoteo: a smartphone application for monitoring and reducing aversive tension in borderline personality disorder patients, a pilot study. Perspect Psychiatr Care 2017 Oct;53(4):289-298. [doi: 10.1111/ppc.12178] [Medline: 27439663]
50. Roncero M, Belloch A, Doron G. A novel approach to challenging OCD related beliefs using a mobile-app: an exploratory study. J Behav Ther Exp Psychiatry 2018 Jun;59:157-160. [doi: 10.1016/j.jbtep.2018.01.008] [Medline: 29425951]
51. Stoll RD, Pina AA, Gary K, Amresh A. Usability of a smartphone application to support the prevention and early intervention of anxiety in youth. Cogn Behav Pract 2017 Nov;24(4):393-404 [FREE Full text] [doi: 10.1016/j.cbpra.2016.11.002] [Medline: 29056845]

52. Shrier LA, Spalding A. 'Just take a moment and breathe and think': young women with depression talk about the development of an ecological momentary intervention to reduce their sexual risk. J Pediatr Adolesc Gynecol 2017 Mar;30(1):116-122. [doi: 10.1016/j.jpag.2016.08.009] [Medline: 27575408]
53. Wen L, Sweeney TE, Welton L, Trockel M, Katznelson L. Encouraging mindfulness in medical house staff via smartphone app: a pilot study. Acad Psychiatry 2017 Oct;41(5):646-650. [doi: 10.1007/s40596-017-0768-3] [Medline: 28795335]
54. Cully J, Teten AL. A Therapist's Guide to Brief Cognitive Behavioral Therapy. MIRECC / CoE Home - Veterans Affairs. 2008. URL: https://www.mirecc.va.gov/visn16/docs/therapists_guide_to_brief_cbtmanual.pdf [accessed 2020-10-22]
55. Snyder M. Self-monitoring process. In: Advances in Experimental Social Psychology. New York, USA: Academic Press; 1979.
56. Snyder M, Gangestad S. On the nature of self-monitoring: matters of assessment, matters of validity. J Pers Soc Psychol 1986 Jul;51(1):125-139. [doi: 10.1037/0022-3514.51.1.125] [Medline: 3735063]
57. Grant AM. Rethinking psychological mindedness: metacognition, self-reflection, and insight. Behav Change 2012 Feb 22;18(1):8-17. [doi: 10.1375/bech.18.1.8]
58. Bennett-Levy J, Thwaites R, Chaddock A, Davis M. Reflective Practice in Cognitive Behavioural Therapy: the Engine of Lifelong Learning. New York, USA: Open University Press; 2001.
59. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIS) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462 [FREE Full text] [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
60. Corrigan PW, Morris SB, Michaels PJ, Rafacz JD, Rüsch N. Challenging the public stigma of mental illness: a meta-analysis of outcome studies. Psychiatr Serv 2012 Oct;63(10):963-973. [doi: 10.1176/appi.ps.201100529] [Medline: 23032675]
61. Neary M, Schueller SM. State of the field of mental health apps. Cogn Behav Pract 2018 Nov;25(4):531-537. [doi: 10.1016/j.cbpra.2018.01.002]
62. Bójko M, Maślankiewicz R. [The situation in Polish child psychiatry at the turn of 2018 and 2019 based on data collected through access to public information]. [Child Abused. Theory, research, practice] 2020;19(1):116-154 [FREE Full text]
63. Baker C. Mental Health Statistics for England: Prevalence, Services and Funding. Research Briefings. 2020. URL: https://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN06988 [accessed 2020-10-22]
64. Grünzig SD, Bengel J, Göhner W, Violetta Krämer L. [low-intensity interventions to reduce depressive symptoms before outpatient psychotherapy - a systematic literature review]. Psychother Psychosom Med Psychol 2019 Jun;69(6):212-223. [doi: 10.1055/a-0630-2397] [Medline: 30045413]
65. Xiang Y, Yang Y, Li W, Zhang L, Zhang Q, Cheung T, et al. Timely mental health care for the 2019 novel coronavirus outbreak is urgently needed. Lancet Psychiatry 2020 Mar;7(3):228-229 [FREE Full text] [doi: 10.1016/S2215-0366(20)30046-8] [Medline: 32032543]
66. Douglas PK, Douglas DB, Harrigan DC, Douglas KM. Preparing for pandemic influenza and its aftermath: mental health issues considered. Int J Emerg Ment Health 2009;11(3):137-144. [Medline: 20437844]
67. Goulia P, Mantas C, Dimitroula D, Mantis D, Hyphantis T. General hospital staff worries, perceived sufficiency of information and associated psychological distress during the A/H1N1 influenza pandemic. BMC Infect Dis 2010 Nov 9;10:322 [FREE Full text] [doi: 10.1186/1471-2334-10-322] [Medline: 21062471]
68. Liu Z, Han B, Jiang R, Huang Y, Ma C, Wen J. Mental health status of doctors and nurses during covid-19 epidemic in China. Soc Sci Netw 2020:-. [doi: 10.2139/ssrn.3551329]
69. Matsuishi K, Kawazoe A, Imai H, Ito A, Mouri K, Kitamura N, et al. Psychological impact of the pandemic (H1N1) 2009 on general hospital workers in Kobe. Psychiatry Clin Neurosci 2012 Jun;66(4):353-360 [FREE Full text] [doi: 10.1111/j.1440-1819.2012.02336.x] [Medline: 22624741]
70. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. JAMA Netw Open 2020 Mar 2;3(3):e203976 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.3976] [Medline: 32202646]
71. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: study design: randomised controlled trials. BJOG 2018 Dec;125(13):1716 [FREE Full text] [doi: 10.1111/1471-0528.15199] [Medline: 29916205]
72. Collins L. Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST). New York, USA: Springer; 2018.
73. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy S. A 'SMART' design for building individualized treatment sequences. Annu Rev Clin Psychol 2012;8:21-48 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032511-143152] [Medline: 22224838]
74. Leonard NR, Casarjian B, Fletcher RR, Praia C, Sherpa D, Kelemen A, et al. Theoretically-based emotion regulation strategies using a mobile app and wearable sensor among homeless adolescent mothers: acceptability and feasibility study. JMIR Pediatr Parent 2018;1(1):e1 [FREE Full text] [doi: 10.2196/pediatrics.9037] [Medline: 30637376]
75. Istepanian RS, Al-Anzi T. m-Health 2.0: new perspectives on mobile health, machine learning and big data analytics. Methods 2018 Dec 1;151:34-40. [doi: 10.1016/j.ymeth.2018.05.015] [Medline: 29890285]

Abbreviations

CBT: cognitive behavioral therapy
EMA: ecological momentary assessment
EMI: ecological momentary intervention
JITAI: just-in-time adaptive intervention
mHealth: mobile health
RCT: randomized controlled trial
VR: virtual reality
WHO: World Health Organization

Edited by G Eysenbach; submitted 04.05.20; peer-reviewed by N Daniëls, H Alhodaib; comments to author 25.07.20; revised version received 25.08.20; accepted 14.09.20; published 12.11.20.

Please cite as:
Marciniak MA, Shanahan L, Rohde J, Schulz A, Wackerhagen C, Kobylińska D, Tuescher O, Binder H, Walter H, Kalisch R, Kleim B
Standalone Smartphone Cognitive Behavioral Therapy–Based Ecological Momentary Interventions to Increase Mental Health: Narrative Review
JMIR Mhealth Uhealth 2020;8(11):e19836
URL: https://mhealth.jmir.org/2020/11/e19836
doi:10.2196/19836
PMID:33180027

Original Paper

Engagement Features in Physical Activity Smartphone Apps: Focus Group Study With Sedentary People

Marco D'Addario[1], PhD; Dario Baretta[1], PhD; Francesco Zanatta[1], MSc; Andrea Greco[2], PhD; Patrizia Steca[1], PhD

[1]Department of Psychology, University of Milan-Bicocca, Milan, Italy

[2]Department of Human and Social Sciences, University of Bergamo, Bergamo, Italy

Corresponding Author:
Francesco Zanatta, MSc
Department of Psychology
University of Milan-Bicocca
Piazza dell'Ateneo Nuovo, 1
Milan, 20126
Italy
Phone: 39 3920317007
Email: francesco.zanatta@unimib.it

Abstract

Background: Engagement with physical activity mobile apps has been reported to be a core precondition for their effectiveness in digital behavior change interventions. However, to date, little attention has been paid to understanding the perspectives, needs, expectations, and experiences of potential users with physical activity mobile apps.

Objective: The aim of this study was to investigate the features that are judged to be important for engagement with a physical activity mobile app and the reasons for their importance.

Methods: A qualitative focus-group methodology with elements of co-design was adopted in this study. Participants reporting sedentary lifestyles and willingness to improve their physical activity behavior through mobile technology were recruited. The focus group sessions consisted of 13 participants (8 men and 5 women, mean [SD] age 41.9 [7.1] years). Two researchers conducted the data analysis independently by using the inductive thematic approach.

Results: Four main themes emerged in relation to the research question and were named as follows: “physical activity participation motives,” “autonomy and self-regulation,” “need for relatedness,” and “smart.” Additionally, 2 subthemes originated from “physical activity participation motives” (ie, “medical guidance” and “weight loss and fitness for health”) and “smart” (ie, “action planning” and “adaptable and tailored”).

Conclusions: Features enhancing autonomy and self-regulation and positively affecting health and physical well-being as well as the need for relatedness, adaptability, and flexibility should be considered as core elements in the engagement of potential users with physical activity mobile apps. The emerged findings may orient future research and interventions aiming to foster engagement of potential users with physical activity apps.

(JMIR Mhealth Uhealth 2020;8(11):e20460) doi:10.2196/20460

KEYWORDS

physical activity; smartphone app; engagement; co-design; thematic analysis; mobile phone

Introduction

Background

Regular physical activity is widely recognized as a protective factor against cardiovascular diseases, diabetes, obesity, and some forms of cancer [1-3]. Nevertheless, a third of the adults worldwide are insufficiently physically active [4]; as a consequence, promoting effective interventions targeting physical activity is a priority. Behavior change interventions that increase physical activity have been shown to be effective [5], and diverse findings supporting their cost-effectiveness have been reported [6]. However, most of these interventions have been delivered with a face-to-face methodology, which is considered, for intrinsic reasons, unsuitable for targeting large populations and supporting behavior change in ecological settings.

Mobile technology provides a valid tool to foster interventions [7,8] and to reach large populations [9]. Nevertheless, little

evidence has been provided so far for the modest effect of web-based and mobile-based interventions on physical activity [9-11]. A recent systematic review and meta-analysis has aimed to determine the role of smartphone apps in increasing objectively measured physical activity in adults and it provided modest evidence supporting short-term effectiveness (eg, up to 3 months) [12]. In this regard, engagement was shown to be a fundamental precondition for digital behavior change interventions (eg, physical activity smartphone apps) and their effectiveness [13]. Nevertheless, digital interventions were shown to have low engagement in different contexts, including controlled trials [10,14] and in relation to health app usage in real life settings [15]. Specifically, the lack of desired features and abandoning health goals were reported as reasons for low engagement and subsequent dropout [16]. Conversely, other studies have suggested higher engagement with health smartphone apps to be potentially supported by interactive features such as social and professional support, self-monitoring, and feedback [17-19]. However, in spite of recommendations [20,21], little attention has been paid to understanding users' perspectives and preferences of these features and implementing them in app functionalities in order to sustain users' engagement with behavior change and health-related goals. Consistently to this point, a recent systematic review unveiled that personal factors and features of the device play a role in influencing participants' motivation to engage with mobile apps promoting physical activity [22]. It was noted that, among the personal factors, prior experience with and rationales for using health apps influenced users' motivation. Moreover, mobile health components were found to potentially facilitate changing users' ways of thinking and self-awareness in relation to physical activity. Indeed, the monitoring features of the apps gave the users opportunities to reflect, which led them to include additional physical activity in their routines. Furthermore, social features, prompts, goal setting, and gamification were noted as determinants of a better experience of mobile health in physical activity, suggesting higher engagement with app use. Thus, ease of use, personalized features, and the possibility to customize the app were shown as relevant factors too. Another recent qualitative study synthesized what is known about influences on the uptake of and engagement with health smartphone apps, suggesting that factors such as physical and psychological capability, physical and social opportunity, and motivation are pivotal and deserve further investigation [23]. Especially for engagement, app instructions, health and well-being information, and visual or numerical summary of progress were considered essential. In addition, self-monitoring, established routines, and safety netting have been taken into account as pivotal features supporting behavioral regulation and consequent engagement with the app.

For these purposes and according to previous observations, a mixed-methods research design with a specific focus on qualitative methodologies (eg, focus groups, interview with open-ended questions, think-aloud studies) may be preferred [21,24,25]. Qualitative methodologies may indeed benefit from various techniques. One in particular that is gaining attention from mobile health researchers is co-design. The term "co-design" refers to the creative collaboration between researchers and end users and to their involvement in the design development process, where the valuable role of the latter relies on their position as "experts of their experience" [26]. *Co-designing*, for instance, a digital intervention in cooperation with potential users encompasses advantages, including better idea generation and better fit between users and the product, higher quality, and more effective products as well as more efficiency in project management [27].

Objective

The aim of this study was to elicit the preferences and perspectives of sedentary potential users toward features of a physical activity mobile app. Hence, we intend to elucidate why specific functionalities and features are judged to be important in the design process of a physical activity mobile app that contributes to increasing potential users' engagement.

Methods

Study Design

This study adopted a focus group methodology with elements of co-design. This study design aimed to address users' perspectives and expectations in relation to a physical activity mobile app. The cocreation of a fit-for-purpose digital product designed with and around the users represents the real added value of this methodology. Co-design activities allow users to express their creativity and contribute to the development process as experts of their own experiences [26]. The project design, procedures, and informed consent form were approved by the ethics committee of the University of Milan-Bicocca.

Study Participants

The participants were recruited according to the following eligibility criteria: (1) age between 30 years and 50 years, (2) no preexisting health conditions that would impede physical activity, (3) no clinical conditions related to physical inactivity (eg, obesity, diabetes), (4) insufficient physical activity screened with the Global Activity Questionnaire [28] and consequently compared to physical activity recommended guidelines (ie, 150 minutes of moderate physical activity or 75 minutes of vigorous physical activity per week), and (5) interest in increasing physical activity behavior through mobile technologies. No monetary compensation for participation was provided.

Sampling

The aim of the focus group and the eligibility criteria were specified in the recruitment materials (social media, snowball sampling methods, and posters placed on the university campus). Recruitment stopped when no further relevant insights of participants' experiences and novel themes emerged from the focus groups [29].

Co-Design Materials

While the materials were being developed, design features were primarily defined in order to fulfil the most relevant physical activity participation motives and to implement the behavior change techniques (eg, action planning, goal setting, problem solving, self-monitoring) [30]. Additional features that are expected to be associated with engagement (ie, rewards, reminders, social support) [18] were also included as stimuli.

A co-design pack was created according to a preliminary revision of the literature. The selected functions and features presented to participants differed between the 2 focus groups. In the first focus group, the features were written on post-its and referred to the behavior change techniques recognized as the most effective in physical activity promotion [5] and to the different motives associated with physical activity [31,32]. In the second focus group, the features were presented as images and referred to those mainly included in the publicly available physical activity apps [33]. Changes of materials between the first and the second focus group were made with the aim to provide the participants with a more engaging and realistic co-design activity, since interacting with images may be more stimulating than post-its. Although the creative activities formed a large proportion of the work, co-design materials and the related creative tasks were exclusively considered as stimuli for participants in focus group conduction and were not included in the data analysis phase.

Procedure

Participants were invited to the University of Milan-Bicocca. Before starting the focus group session, participants were given the information sheet describing the nature of the study. They subsequently signed the informed consent form. As the first focus group activity, participants introduced themselves and shared their previous personal experiences with physical activity and with mobile apps. Subsequently, a folder with materials for co-design was provided. Participants were asked to design their own physical activity app according to their preferences for specific functionalities and features. According to the user-centered methodology and the iterative nature of the co-design process, materials and the track for the following co-design sessions were consequently adapted in order to better elicit users' views about what specific functionalities are expected to be important and why. The first session consisted of evaluating and designing potential physical activity app features (functionalities and features were written on post-its given to participants at the beginning of the focus group session) according to different topics (ie, feedback and monitoring, goal setting and planning, challenges and social features, problem solving). A brief discussion was then conducted when each topic was completed. In the latter session, participants were asked to cut out images of features printed on a paper sheet and paste them on a printed smartphone frame according to their preferences. Once the whole activity was accomplished, a general discussion was conducted. In both co-design sessions, participants were also provided with blank post-its to suggest any further functionalities they would like to see implemented in a physical activity app. Changes to the co-design task relied on the aim to characterize the activity with more concrete stimuli and to stress the creative act of cocreating the product. Therefore, this did not influence the scope of focus group sessions (ie, elicit potential users' views and preferences for design features) but, rather, provided an opportunity for a better expression of participants' views and perspectives.

Data Analysis

The focus group sessions were audio-recorded and transcribed verbatim. The transcripts were qualitatively analyzed using inductive thematic analysis [34]. Thematic analysis consists of 6 phases: (1) familiarizing with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report. Given the exploratory nature of the thematic analysis methodology, it was expected that diverse themes with different contents might emerge [34]. For this reason, the themes might not be strictly connected to engagement. Therefore, any theme that emerged and that was not necessarily related to engagement was accepted as an informative finding. Data and repeated patterns that were considered pertinent to the aim of the study were coded by 2 researchers working independently. New inductive codes were labelled as they were identified during the coding process and the results of the coding were iteratively revised. The next stage involved searching for themes; both researchers reviewed their own generated codes one by one, organizing the findings in order to combine the different codes that have been considered focusing on the same aspect. Finally, the codes generated by the 2 researchers were compared and subsequently organized into themes. Discussion within the entire research group was conducted in order to reach a consensus on the final themes. Once themes were defined and named, examples of transcripts were selected to corroborate themes on the basis of their representativeness and relevance. Data were analyzed in their original language in order to preserve the participants' original meanings, while coding and themes were formulated in English only. The selected examples of transcripts were translated into English for illustration purposes.

Results

Overview of the Findings

Two focus groups were formed (N=13, 8 men and 5 women). The mean (SD) age of the participants was 41.9 (7.1) years. All the participants reported doing less than 120 minutes of light-to-moderate physical activity per week; 4 participants reported having previously used smartphone apps to support physical activity.

Four themes were developed in relation to the research question and were named as follows: "physical activity participation motives," "autonomy and self-regulation," "need for relatedness," and "smart." Two subthemes were developed in relation to the "physical activity participation motives" theme: "medical guidance" and "weight loss and fitness." Additionally, 2 subthemes originated from the "smart" theme: "action planning" and "adaptable and tailored" (see Table 1 for a brief description of the themes and subthemes).

Table 1. Summary of the themes and subthemes that emerged from the thematic analysis.

Themes and subthemes	Description
Physical activity participation motives	The intention to do physical activity is more related to health and well-being or duty motives, rather than intrinsic motivation or pleasure.
Medical guidance	A physical activity app should contain features that include medical support, providing feedback about physiological parameters related to health.
Weight loss and fitness for health	A physical activity app should contain features that contribute to physical benefits and cardiovascular fitness.
Autonomy and self-regulation	Physical activity app design should include features (eg, monitoring, goal setting, feedback) that contribute to increasing self-regulation and autonomous behaviors.
Need for relatedness	It is crucial that a physical activity app contains features that leverage the motivational component of social support and promote privacy, safety, and comparison with oneself.
Smart	Subjects ask for a flexible physical activity mobile app that provides context-aware suggestions and goal-oriented feedback.
Action planning	Physical activity app design should include activity suggestions in relation to users' personal goals and flexible programs.
Adaptable and tailored	Potential users would prefer a customized physical activity mobile app that can deliver tailored suggestions and that is built on their own characteristics and motivations.

Physical Activity Participation Motives

Participants highlighted the importance of using a physical activity app focusing on participation motives. Most of them reported that their intention to practice physical activity did not originate from intrinsic motivation or pleasure, but from health and well-being–related or duty motives.

> *…I consider physical activity as a responsibility to feel healthy, rather than a pleasure. It is an effort that I make to feel better.* [Male #1]

> *…For some people, fitness is a pleasure. Personally, I would exercise because it would make me healthy, not because I'd like it.* [Male #2]

Specifically, physical activity participation motives can be attributable to the desire for medical guidance and a specific focus on weight loss and fitness.

Subtheme 1 of Physical Activity Participation Motives: Medical Guidance

Positive opinions about the opportunity to be informed and supported by medical sources (eg, medical doctors or devices) were expressed. This aspect would contribute to making the app more credible, reliable, and trustworthy. Moreover, thanks to the presence of a medical perspective, participants expected the app to be able to provide feedback about the physiological parameters strictly related to health.

> *…Physical activity should be supported and monitored by a sports doctor, and after an exercise electrocardiogram… It could be dangerous otherwise!* [Male #3]

> *…I would like to be supported by a virtual personal trainer, but I need something more such as a doctor, an expert who gives reliable and evidence-based recommendations.* [Male #4]

Subtheme 2 of Physical Activity Participation Motives: Weight Loss and Fitness for Health

Focus group participants remarked the importance of physical activity for weight loss and fitness, thereby contributing to physical benefits and cardiovascular health.

> *…My eating habits are important but my physical well-being is also a fundamental aim for me: to get my fitness back, to control my body weight and my heart […] I always imagine an app matching with my main aim, which is weight loss, blood pressure and heart control. I imagine an app as an instrument supporting physical well-being.* [Male #5]

> *…I am interested because my hope is to lose weight. I have perennial fight with my weight…"* [Female #2]

Autonomy and Self-Regulation

Most of the participants reported monitoring, goal setting, and feedback as the key components for a physical activity app design. These features would help increase not only self-regulation in physical activity but also autonomous behaviors.

> *…It must be something [app] that helps me to reach my goal. It should be a calculator of what I did and what I should have done…* [Male #5]

> *…I consider information like how far and how long I have run, the distance and other similar data that I consider the starting point. These may then help to add further elements that an app should offer.* [Female #1]

> *…you receive a notification on your smartphone like "you are worsening, your conditions are decreasing"* [Male #1]

In particular, regarding monitoring and feedback features, most participants said that they were interested in having records and statistics about their itineraries.

> *…I would like my physical activity to be monitored and analyzed with charts and stats. I would prefer to see daily, weekly, and monthly statistics.* [Male #2]
>
> *…I would like the app to track my paths. It could motivate me to do more, improving my results. Also, I imagine an app providing an archive of previous paths.* [Male #4]

Need for Relatedness

Focus group participants expressed social support as a fundamental component that could help to create higher levels of daily commitment, to make exercise a more enjoyable activity, as well as to increase motivation and overcome laziness.

> *…I should find someone who motivates my goal setting. Not daily goals, but goals helping me to use the app regularly […] It would push me to share with others, to be more active…* [Female #2]
>
> *…If you have an appointment, then you can't quit.* [Female #3]
>
> *…I have always been very lazy. Laziness leads me to be idle when I'm alone. So, it would be better if there was someone with me.* [Female #4]
>
> *…You have to find someone at your own pace.* [Male #2]

Although peer support was shown to be a core element, most of the participants expressed avoidance to social comparison, exhibitionism, and competition, and preferred privacy, safety, and comparison with oneself.

> *…I almost never share my private life […] I don't think it is interesting to know or to let others know what I usually do.* [Female #1]
>
> *…It could be dangerous to let you know what I do, or to let me know what you do every day.* [Male #3]
>
> *…I would exercise for personal motives and not for competing […] Everything I do is for a comparison with myself: I like seeing my own improvements.* [Male #1]

Furthermore, the possibility to interact with a virtual personal trainer who is aware of the personal motivation, abilities, and preferences was reported as a core element.

> *…I would like to rub the smartphone to bring up a personal trainer […] I do not perceive the human component. I would need the relationship with a personal trainer, with someone able to show the exercises right in front of you.* [Female #5]
>
> *…I would need an app like a personal trainer who monitors me […] Someone who supports me.* [Female #2]
>
> *…It should be something that gives tailored hints, knowing one's motivations… like a parent who knows you, who loves you, who says what to do.* [Male #3]
>
> *…and then, I would like something like Siri that asks to you before you start: "How do you feel today?" "Today, I don't want to put in too much effort" "Ok, I will choose this path for you, then!* [Female #3]

Smart

Participants imagined a physical activity mobile app presenting flexibility and adaptability toward individual circumstances and motivations. For this purpose, the app should provide context-aware suggestions, goal-oriented feedback, and action-oriented planning for a tailored support in physical activity.

Subtheme 1 of Smart: Action Planning

Participants expressed the preference to use a physical activity app that consisted of features that suggest activities in relation to personal goals and flexible programs.

> *…I would like the app to propose a training. An app suggesting a training according to my goal.* [Male #6]
>
> *…an app helping me to identify a lighter training instead of forcing me when I am particularly tired…* [Female #3]
>
> *…I would need an app suggesting a path, for example, "Today, you have one hour available and you could choose this path. There is a park nearby.*" [Female #2]
>
> *…an app identifying paths or outdoor gyms. So, for example, when I don't know where to go to run it could tell me, "These are some paths you might be interested in.*" [Female #1]

Subtheme 2 of Smart: Adaptable and Tailored

Most of the participants believed that a physical activity app should be customized to users' characteristics and motivations and should also be able to read and detect behaviors and needs and deliver tailored suggestions.

> *…to customize it as much as possible, it should be tailored […] If the app knows that you have high blood pressure, it will not suggest you go running for 3 hours.* [Female #1]
>
> *…It should be as complete as possible. It should not be based on only one parameter, but rather on different characteristics […] It should be customized, for example, according to one's psychological and physical condition at the specific moment of the day.* [Female #2]
>
> *…Based on the data on your goals you fill in when registering, the app should modulate the difficulty and give some short term results to achieve your goals.* [Male #7]
>
> *…Everything needs to be flexible considering that I am flexible.* [Male #3]

Discussion

Principal Results

In this study, we investigated the potential features of a physical activity app that are judged to be important for engagement by sedentary potential users and the reasons for their importance. A focus group methodology with elements of co-design was adopted in order to involve participants who reflected about the app features and researchers who characterized the design process. The findings of this study suggest that features that enhance users' autonomy and self-regulation and those that focus on the impact physical activity has on health and physical well-being were considered to be relevant for engagement. The need for relatedness was considered as a trigger for exercise if it adopts a supportive and motivational style. For this purpose, both a virtual personal trainer with human qualities embedded into the app and opportunities to develop a network of physical activity peers were judged as relevant features. A different aspect noticed in this study was that participants seemed to give importance to the private dimension of exercising and, thus, expressed a preference for self-comparison, privacy, and safety rather than competition, exhibitionism, and social comparison. Finally, participants believed that a smart, tailored interaction between user and app would foster a more effective engagement by providing an adaptable flexible approach for users to achieve personal goals.

Comparison With Prior Work

The intentions and motives related to exercise seemed to be mostly oriented toward health and physical well-being concerns, which were partially coherent with those reported in a previous research [19]. In the previous study [19], specific features focusing on fitness, nutrition, and weigh loss were listed as the main reasons for engaging in physical activity. In addition, a preference for particular suggestions and the possibility to track relevant health-related parameters was indicated. As for what emerged in this research, improving cardiovascular fitness and weight loss and monitoring of health-related indicators were repeatedly reported as the primary motivations for being more active and healthier. To address this request, physical activity apps should take into account a wider and more integrated approach for physical activity by involving medical guidance to help users to achieve and monitor their health-related goals (eg, weight control, aerobic improvements). For these purposes, new emerging mobile technology such as self-tracking devices may provide more reliable, integrated information about physical activity behavior and its outcomes (eg, indicators of weight loss, index of fitness, heart rate).

The preferences of the potential users for autonomy and self-regulation features were consistent with those reported in a previous research [19,35] and corroborated the results from a recent review [18], suggesting that behavior change techniques such as goal setting, feedback, and self-monitoring are associated with higher engagement. Specifically, these features allow participants to monitor their improvements, highlighting any potential discrepancy between their physical activity behavior and their goal. Moreover, self-regulation features (ie, goal setting, behavior monitoring, receiving feedback) have been shown to be effective in increasing physical activity [36]. Similar convergent empirical evidence has clearly validated the idea that considering self-regulation features in a physical activity app allows users to self-organize their experience and behavior, thereby fostering their autonomy in exercising [19]. Furthermore, they deserve to be considered as the core component of physical activity apps, as they have already been expected to increase engagement [17,18], be appreciated by users [37], and be associated with physical activity intervention effectiveness [5,38]. It is noteworthy that although commercial physical activity apps often implement such features [39], this might be done in a more theory-driven way as highlighted for goal setting in a recent review and content analysis of commercial apps [40].

The need for relatedness was perceived by potential physical activity app users as a core element for increasing their engagement with the app. Previous research has shown social support features not to be as relevant as the self-regulation feature [41]. In contrast, a recent study [19] supported the motivational aspect of social interactions as an element supporting and fostering higher levels of commitment. Indeed, it was unveiled that although varying types and levels of needs for relatedness with others were expressed, connection with peers and coaching features embedded into the app play a promoting role, thereby increasing motivation and making physical activity more enjoyable. Consistently, the findings of this study show that human interactions with peers and virtual personal trainers are judged to be a trigger for being more active and overcoming laziness, as well as a fundamental feature in support and consequent dedication. Indeed, features connecting the user with a virtual coach or a group of people with similar goals were expected to be important for engagement because of the formation of a shared commitment and the opportunity to be emotionally supported. The findings in this study are consistent with those reported in a previous study [42], suggesting that working alliance and the desire to continue using a digital behavior change intervention to promote physical activity are increased by the support provided by a computer interface with human-relational skills (eg, empathy, social dialogue). Furthermore, a scoping review of web-based interventions showed supportive virtual coaches to be a potentially valuable remedy for low adherence to digital interventions [43]. In accordance with previous research targeting different behaviors [18,37], potential users of physical activity apps are reluctant to share information with social networks about their physical activity. Indeed, users perceived physical activity behavior change as a personal path and did not see merit in showing it to other users. This avoidance of competition and social comparison is consistent with that reported in a previous study [19], as competence and competition were listed among the less relevant motives for practicing physical activity and social comparison was revealed to be one of the less-liked behavior change techniques. Specifically, the latter was also considered as an obstacle for motivation to exercise due to potential exposure to others' negative judgments. Furthermore, participants definitely rejected the idea of thinking about physical activity from a competitive perspective, probably because any source of social comparison

or competition with more successful users might discourage them, thereby constituting an additional stressful element.

Finally, participants imagined a smart, tailored physical activity app that is customized to "understand" the users' preferences and expectations and consequently support them as and when required. For example, action planning strategies based on the user's timetable and suggestions tailored to the user's level of progress toward a specific goal were expected to be more engaging and effective. The findings in this study corroborate previous observations, indicating that providing tailored physical activity plans based on personal goals and current progress would help users implement their engagement and intentions to exercise [19]. A similar need for a tailored approach was also found in a previous research investigating users' preferences for design features related to smartphone apps for drinking reduction and smoking cessation [18]. Additionally, participants emphasized their interest in a flexible physical activity app that adapts the intervention content to the specific context and variations in the users' motivational and emotional states in real time. This suggestion is in line with recent emerging research that has focused on the development of just-in-time adaptive interventions [44] and tailoring physical activity interventions on motivational aspects such as self-efficacy beliefs [45,46].

Limitations

The limitations of this study include the design of the study: a nonlongitudinal design could have limited the long-term investigation of the influence of the features emerged in users' engagement. For this purpose, further qualitative studies may be conducted comprising a follow-up aimed to evaluate the possible effects on users' engagement after a period of tailored smartphone app usage. Another limitation of the study is the narrow age range of participants. Actually, the decision to only include adults aged 30-50 years was based on the belief that this would have improved the quality of the findings. Perski et al [18] reported that engagement with digital behavior change interventions is influenced by age (showing a trend toward a positive association between engagement and older age). However, as the moderating role of age on engagement cannot be easily controlled for in qualitative studies, we preferred to avoid recruiting participants younger than 30 years and older than 50 years in order to prevent collecting data from a heterogeneous sample with varying motivations and subjective experiences. Further investigations may be helpful to explore the role of age. Furthermore, the nature of the focus group may have limited the expression of the participants' true attitudes because of social desirability bias. Finally, as there can be a difference between what is expected to help engagement and what actually has an influence on engagement, a prospective study including a period of physical activity smartphone app use is recommended. In this way, experiencing the app may lead to understanding whether the features actually have the same impact on engagement as expected.

Conclusions

The findings in this study provide informative data concerning sedentary potential users' preferences and perspectives toward physical activity smartphone app features. Consistent with that reported in prior works, improving health and physical well-being were indicated as the primary motivations for exercising, as was the possibility to track relevant health-related parameters with the additional support of medical sources. Similarly, features strictly related to the behavior change techniques (ie, goal setting, feedback, and self-monitoring) were preferred to support autonomy and self-regulation. Furthermore, the preference of tailored and customized features "understanding" user's expectations and level of progress was shown and this corroborated prior observations. Finally, features that leverage the motivational component of social support and that promote privacy, safety, and self-comparison were shown to be crucial. Nevertheless, they still represent an open question, as contrasting results emerged in prior studies. In order to increase users' engagement with physical activity smartphone apps and, thus, to improve their effectiveness, the features described here should be taken into account in future design processes and in future research aiming to broaden knowledge on mobile health in relation to physical activity.

Authors' Contributions

DB conducted the focus group sessions. DB and MD performed the thematic analysis and the entire research group was involved in reaching a consensus on the final themes. All the authors contributed to paper writing and revision.

Conflicts of Interest

None declared.

References

1. Arem H, Moore SC, Patel A, Hartge P, Berrington de Gonzalez A, Visvanathan K, et al. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. JAMA Intern Med 2015 Jun;175(6):959-967 [FREE Full text] [doi: 10.1001/jamainternmed.2015.0533] [Medline: 25844730]
2. Fogelholm M. Physical activity, fitness and fatness: relations to mortality, morbidity and disease risk factors. A systematic review. Obes Rev 2010 Mar;11(3):202-221. [doi: 10.1111/j.1467-789X.2009.00653.x] [Medline: 19744231]
3. World Health Organization. Global recommendations on physical activity for health. In: WHO Library Cataloguing-in-Publication Data. Geneva: World Health Organization; 2010.
4. World Health Organization. Prevalence of insufficient physical activity. Global Health Observatory (GHO) data. URL: https://www.who.int/gho/ncd/risk_factors/physical_activity_text/en/ [accessed 2020-04-09]

5. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011 Feb 18;11:119. [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
6. Abu-Omar K, Rütten A, Burlacu I, Schätzlein V, Messing S, Suhrcke M. The cost-effectiveness of physical activity interventions: A systematic review of reviews. Prev Med Rep 2017 Dec;8:72-78 [FREE Full text] [doi: 10.1016/j.pmedr.2017.08.006] [Medline: 28856084]
7. Pagoto S, Bennett GG. How behavioral science can advance digital health. Transl Behav Med 2013 Sep;3(3):271-276 [FREE Full text] [doi: 10.1007/s13142-013-0234-z] [Medline: 24073178]
8. Taj F, Klein MCA, van Halteren A. Digital Health Behavior Change Technology: Bibliometric and Scoping Review of Two Decades of Research. JMIR Mhealth Uhealth 2019 Dec 13;7(12):e13311 [FREE Full text] [doi: 10.2196/13311] [Medline: 31833836]
9. Müller AM, Alley S, Schoeppe S, Vandelanotte C. Int J Behav Nutr Phys Act 2016 Oct 10;13(1):109 [FREE Full text] [doi: 10.1186/s12966-016-0434-2] [Medline: 27724911]
10. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012 Apr 30;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
11. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res 2012 Nov 21;14(6):e161 [FREE Full text] [doi: 10.2196/jmir.2171] [Medline: 23171838]
12. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can Smartphone Apps Increase Physical Activity? Systematic Review and Meta-Analysis. J Med Internet Res 2019 Mar 19;21(3):e12053 [FREE Full text] [doi: 10.2196/12053] [Medline: 30888321]
13. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
14. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
15. Consumer Health Information Corporation. Motivating Patients to Use Smartphone Health Apps. URL: http://www.prweb.com/releases/2011/04/prweb5268884.htm [accessed 2020-04-09]
16. Murnane E, Huffaker D, Kossinets G. Mobile health apps: adoption, adherence, and abandonment. In: UbiComp/ISWC'15 Adjunct: Proceedings of the 2015 of the 2015 ACM International Symposium on Wearable Computers. 2015 Sep 14 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings; 2015; Osaka Japan p. 261-264. [doi: 10.1145/2800835.2800943]
17. 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]
18. 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]
19. 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 Feb 08;7(2):e11636 [FREE Full text] [doi: 10.2196/11636] [Medline: 30735143]
20. 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]
21. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
22. Carter DD, Robinson K, Forbes J, Hayes S. Experiences of mobile health in promoting physical activity: A qualitative systematic review and meta-ethnography. PLoS One 2018;13(12):e0208759 [FREE Full text] [doi: 10.1371/journal.pone.0208759] [Medline: 30557396]
23. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the Uptake of and Engagement With Health and Well-Being Smartphone Apps: Systematic Review. J Med Internet Res 2020 May 29;22(5):e17572 [FREE Full text] [doi: 10.2196/17572] [Medline: 32348255]
24. Van Gemert-Pijnen JEWC, Peters O, Ossebaard HC. Improving eHealth. Netherlands: The Hague, The Netherlands: Eleven International Publishers; 2013.
25. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: A scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]

26. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068]
27. Steen M, Manschot M, De Koning N. Benefits of co-design in service design projects. Int J Des. 2011. URL: http://www.ijdesign.org/index.php/IJDesign/article/view/890/346 [accessed 2020-01-03]
28. Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). J Public Health 2006 Mar 2;14(2):66-70. [doi: 10.1007/s10389-006-0024-x]
29. Guest G, Bunce A, Johnson L. How Many Interviews Are Enough? Field Methods 2016 Jul 21;18(1):59-82. [doi: 10.1177/1525822X05279903]
30. 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]
31. Frederick CM, Ryan RM. Differences in motivation for sport and exercise and their relationships with participation and mental health. J Sport Behav. 1993. URL: https://selfdeterminationtheory.org/SDT/documents/1993_FrederickRyan_Differencesinmotivation.pdf [accessed 2020-03-01]
32. Ryan RM, Frederick CM, Lepes D, Rubio N, Sheldon KM. Intrinsic motivation and exercise adherence. Int J Sport Psychol. 1997. URL: https://selfdeterminationtheory.org/SDT/documents/1997_RyanFrederickLepesRubioSheldon.pdf [accessed 2020-02-01]
33. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of Publicly Available Physical Activity Apps: Review and Content Analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
34. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
35. Middelweerd A, van der Laan DM, van Stralen MM, Mollee JS, Stuij M, te Velde SJ, et al. What features do Dutch university students prefer in a smartphone application for promotion of physical activity? A qualitative approach. Int J Behav Nutr Phys Act 2015 Mar 01;12:31 [FREE Full text] [doi: 10.1186/s12966-015-0189-1] [Medline: 25889577]
36. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: A scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
37. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
38. Knittle K, Nurmi J, Crutzen R, Hankonen N, Beattie M, Dombrowski SU. How can interventions increase motivation for physical activity? A systematic review and meta-analysis. Health Psychol Rev 2018 Sep;12(3):211-230. [doi: 10.1080/17437199.2018.1435299] [Medline: 29385950]
39. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
40. Baretta D, Bondaronek P, Direito A, Steca P. Implementation of the goal-setting components in popular physical activity apps: Review and content analysis. Digit Health 2019;5:2055207619862706 [FREE Full text] [doi: 10.1177/2055207619862706] [Medline: 31360535]
41. Belmon LS, Middelweerd A, Te Velde SJ, Brug J. Dutch Young Adults Ratings of Behavior Change Techniques Applied in Mobile Phone Apps to Promote Physical Activity: A Cross-Sectional Survey. JMIR Mhealth Uhealth 2015 Nov 12;3(4):e103 [FREE Full text] [doi: 10.2196/mhealth.4383] [Medline: 26563744]
42. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]
43. Scholten M, Kelders SM, van Gemert-Pijnen JE. A Scoped Review of the Potential for Supportive Virtual Coaches as Adjuncts to Self-guided Web-Based Interventions. : Springer; 2017 Presented at: 12th International Conference, PERSUASIVE 2017; April 4-6, 2017; Amsterdam, The Netherlands p. 43-54. [doi: 10.1007/978-3-319-55134-0_4]
44. Hekler EB, Michie S, Pavel M, Rivera DE, Collins LM, Jimison HB, et al. Advancing Models and Theories for Digital Behavior Change Interventions. Am J Prev Med 2016 Nov;51(5):825-832 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.013] [Medline: 27745682]
45. Pirolli P. A computational cognitive model of self-efficacy and daily adherence in mHealth. Transl Behav Med 2016 Dec;6(4):496-508 [FREE Full text] [doi: 10.1007/s13142-016-0391-y] [Medline: 27848213]
46. Baretta D, Sartori F, Greco A, D'Addario M, Melen R, Steca P. Improving Physical Activity mHealth Interventions: Development of a Computational Model of Self-Efficacy Theory to Define Adaptive Goals for Exercise Promotion. Advances in Human-Computer Interaction 2019 Mar 04;2019:1-11. [doi: 10.1155/2019/3068748]

Edited by G Eysenbach, L Buis; submitted 19.05.20; peer-reviewed by M Klein, L McCann, D Szinay; comments to author 03.07.20; revised version received 18.08.20; accepted 15.10.20; published 16.11.20.

Please cite as:
D'Addario M, Baretta D, Zanatta F, Greco A, Steca P
Engagement Features in Physical Activity Smartphone Apps: Focus Group Study With Sedentary People
JMIR Mhealth Uhealth 2020;8(11):e20460
URL: http://mhealth.jmir.org/2020/11/e20460/
doi:10.2196/20460
PMID:33196450

Original Paper

Impact of Pediatric Mobile Game Play on Healthy Eating Behavior: Randomized Controlled Trial

Yi-Chin Kato-Lin[1], PhD; Uttara Bharath Kumar[2], MHS; Bhargav Sri Prakash[3], MEng; Bhairavi Prakash[4], MSc; Vasini Varadan[5], MEcon; Sanjeeta Agnihotri[6], MSW; Nrutya Subramanyam[7], MBBS, PGDFM, MRCGP(INT); Pradeep Krishnatray[6], PhD; Rema Padman[8], PhD

[1]Hofstra University, Hempstead, NY, United States

[2]Center for Communication Programs, Johns Hopkins University, Baltimore, MD, United States

[3]FriendsLearn Inc, Palo Alto, CA, United States

[4]The Mithra Trust, Chennai, India

[5]Mind in Motion, Chennai, India

[6]Center for Communication and Change – India, New Delhi, India

[7]Seethapathy Clinic & Hospital, Chennai, India

[8]The Heinz College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, PA, United States

Corresponding Author:
Rema Padman, PhD
The Heinz College of Information Systems and Public Policy
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA,
United States
Phone: 1 412 268 2180
Email: rpadman@cmu.edu

Abstract

Background: Video and mobile games have been shown to have a positive impact on behavior change in children. However, the potential impact of game play patterns on outcomes of interest are yet to be understood, especially for games with implicit learning components.

Objective: This study investigates the immediate impact of fooya!, a pediatric dietary mobile game with implicit learning components, on food choices. It also quantifies children's heterogeneous game play patterns using game telemetry and determines the effects of these patterns on players' food choices.

Methods: We analyzed data from a randomized controlled trial (RCT) involving 104 children, aged 10 to 11 years, randomly assigned to the treatment group (played fooya!, a dietary mobile game developed by one of the authors) or the control group (played Uno, a board game without dietary education). Children played the game for 20 minutes each in two sessions. After playing the game in each session, the children were asked to choose 2 out of 6 food items (3 healthy and 3 unhealthy choices). The number of healthy choices in both sessions was used as the major outcome. We first compared the choice and identification of healthy foods between treatment and control groups using statistical tests. Next, using game telemetry, we determined the variability in game play patterns by quantifying game play measures and modeled the process of game playing at any level across all students as a Markov chain. Finally, correlation tests and regression models were used to establish the relationship between game play measures and actual food choices.

Results: We saw a significant main effect of the mobile game on number of healthy foods actually chosen (treatment 2.48, control 1.10; $P<.001$; Cohen $d=1.25$) and identified (treatment 7.3, control 6.94; $P=.048$; Cohen $d=.25$). A large variation was observed in children's game play patterns. Children played an average of 15 game levels in 2 sessions, with a range of 2 to 23 levels. The greatest variation was noted in the proportion of scoring activities that were highly rewarded, with an average of 0.17, ranging from 0.003 to 0.98. Healthy food choice was negatively associated with the number of unhealthy food facts that children read in the game (Kendall $\tau=-.32$, $P=.04$), even after controlling for baseline food preference.

Conclusions: A mobile video game embedded with implicit learning components showed a strong positive impact on children's food choices immediately following the game. Game telemetry captured children's different play patterns and was associated

with behavioral outcomes. These results have implications for the design and use of mobile games as an intervention to improve health behaviors, such as the display of unhealthy food facts during game play. Longitudinal RCTs are needed to assess long-term impact.

Trial Registration: ClinicalTrials.gov NCT04082195; https://clinicaltrials.gov/ct2/show/NCT04082195, registered retrospectively.

(JMIR Mhealth Uhealth 2020;8(11):e15717) doi:10.2196/15717

KEYWORDS

pediatric obesity; mobile games; implicit learning; healthy eating behavior evaluation; game telemetry analysis

Introduction

Background and Significance

Obesity is an increasingly common epidemic in children, with serious long-term health consequences and high health care costs [1,2]. While there are many factors that contribute to overweight and obesity, dietary decisions are a leading cause [1]. Threats to pediatric health risk from the double burden of overweight and malnutrition or undernutrition heightens the need to influence children's dietary lifestyle habit formation through evidence-based, scalable, digital therapeutic methods [3]. There is a clear need to identify effective methods for improving dietary intake and physical activity habits early in life that children find engaging and can remain intrinsically motivated to adopt.

Several diet-related video games have been designed and evaluated for children in recent years, such as Diab [4], Squire's Quest! [5], Squire's Quest! II [6], Quest to Lava Mountain [7], D.W.'s Unicorn Adventure [8], Alien Health Game [9], Fitter Critters [10], Creature-101 [11], Virtual Sprouts [12], Fit, Food, Fun [13], and Garfield vs Hotdog [14]. Most of the video games being evaluated predominantly use explicit education strategies such as providing answers, feedback, instructions, or suggestions to the players [4-6,9-13]. Implicit learning is another strategy that educates players without making them aware of the fact [15]. This is a rapidly growing field due to the high popularity of video games on mobile devices in general and serious games for health and education in particular. Lumosity [16] is one such well-known game, which aims to improve behavioral and cognitive performance by stimulating players' brain function [17]. Several studies have shown its positive effect [17,18]. Without explicitly educating the players, implicit learning embedded in video games has also been found to improve actual sports behavior and physics knowledge [19,20]. However, evidence of the effects of games with implicit education strategies on pediatric healthy eating behavior is still very limited.

Serious games are generally considered an effective intervention for influencing healthy eating behaviors in children [21,22], with some exceptions [9,12,14]. There are studies providing evidence of their positive effects on food intake [4] and healthy eating attitude and self-efficacy [10]. Hence, personalizable gamification and learnification on mobile devices may be one approach for children to learn about healthy dietary habits in a fun and enjoyable way, especially mobile games embedded with implicit learning strategies. Video games include many levels of challenges, imaginative virtual worlds, and the opportunity to navigate them in distinct ways. While much is known about the design of serious games, recent research has highlighted the gap in understanding probable links between game-playing behaviors and observed outcomes so that game design and redesign can be informed through evidence-based knowledge and practices [23-25]. Existing randomized controlled trials (RCTs) predominantly compare health-related outcomes between treatment and control groups without examining how the games are played [21,22,26]. To design appropriate interventions in the game environment for children's health-related behavior formation and change, we need to learn more about the underlying patterns of player behaviors or engagement evidenced during gameplay.

The availability of highly granular game telemetry—data obtained about the actual clicks made by players as they navigate the game—provides a unique opportunity to understand the potential associations between game-playing behaviors and observed outcomes. These clickstream data are usually used to identify flaws in game design and improve user retention measures such as churn and attrition [27,28]. The growing field of game analytics mostly focuses on using game telemetry to discover and communicate meaningful characteristics in the context of game development and impact [29]. For example, Westera and colleagues [30] use aggregated game metrics, such as total time played, number of user actions, and number of videos viewed, to describe student engagement and associate these aggregated variables with student learning outcomes. Similarly, Sharma and colleagues [7] examine the relations between aggregate game metrics (time spent in game play and levels played) and health behavioral outcomes. However, the dynamic of game play was not captured, missing potential insights from the flow of the game play, resulting in a gap in existing studies [31]. Additionally, identifying which paths result in the longest game time and what game features engage players in these paths require methods to learn paths from detailed game telemetry and examine the features that define these paths in unique ways. Loh and Sheng [32] explore the dynamics of game play and portray players' game play behaviors by constructing game play paths/sequences for each user and using the paths to classify game play experts from novices. However, in the current literature, limited attention has been paid to the impact of these paths, which incorporate aggregate game play data and dynamics of the game play, on health-related behavioral outcomes.

In this study, we analyze the relationship between game play patterns and dietary health–related behavioral outcomes using

the mobile serious game fooya!, an action video game designed to promote healthy eating and physical activity in children [33]. The primary education components are implicitly embedded in the game, aiming to make players think deeply and strategically, without the game explicitly showing the right answers. For example, the avatar in the game collects highly prized coins by fighting against enemies made up of unhealthy foods. Building on prior descriptive work [34], we first examine the immediate impact of fooya! on children's food choices in comparison with a game without educational components (Uno board game) using an RCT study design. We then examine the various patterns of player engagement and investigate their effects on player food choices. Observed food choice data, game telemetry, and survey data are collected from the RCT, and a relevant subset of the data are analyzed in this study.

This study adds to the literature on serious health games for children by evaluating the effect of a video game with implicit learning on dietary behavior. In addition, this study makes two novel contributions to the literature. First, while it is well known that children have different play patterns during game play, our research quantifies and measures these differences by looking into the detailed game telemetry and constructing measurable variables such as the length of the activity sequence, number of transitions between actions, and frequency of reading food facts in contrast with analyzing survey, interview, and observational data and serves as the first step for evidence-based personalization. Second, this paper provides novel and granular insights into designing more impactful interventions on mobile games using the lens of examining the complex interactions between game playing patterns and health behaviors.

Objectives

The objectives of this study are to (1) examine the immediate impact of a pediatric dietary mobile game with implicit learning on children's actual food choices, (2) quantify children's heterogeneous game play patterns, and (3) understand the effects of game engagement by associating game play patterns with players' actual food choices.

Theoretical Background

The key knowledge for eating healthfully is implicitly embedded in fooya! (eg, the fact that the avatar's speed and body shape vary in response to the type of food intake) and also presented as nutrition facts at the end of each level. The output of implicit learning, representing the practical knowledge needed to perform a behavior [15], is necessary for behavioral change [35]. This knowledge can be delivered effectively in a game if its engaging nature can trigger children's intrinsic motivation to pay attention to and engage in the game-like neurocognitive training [36]. This enactive learning process—learning by doing and experiencing the consequences of one's actions—that the players experience virtually in the game, through the actions of the avatar, also contributes to behavioral change [35,37]. As promoted by the Institute for the Learning Sciences [38], learning by doing is an approach for mastering some real-world tasks. The learning experience in the game can be further transferred to real-world principles, without instructions to concentrate on problem similarity [39,40]. This suggests that actions that result in healthier body shape are retained and those that lead to negative consequences, such as heavier body shape, will be refined or discarded in the real world, which motivates our objective 1.

While children may learn from the experience, they may learn in different ways. Learning styles are often used to describe the distinct characteristics of children during their learning process, such as their use of problem-solving strategies and decision-making behaviors [41-44]. Different learning styles are found to be related to children's learning outcomes, such as grades, spoken language, and general conduct [41,43,45-47]. The form-function shifts model by Saxe [48,49] assumes that development of thinking can be understood through examination of children's goal-directed activity, which is also true in game environments. For example, Buckley and Doyle [47] find several relationships between learning styles and game engagement. Hassan and colleagues [42] are able to predict students' learning styles using play logs with 72% to 78% accuracy. Nasir [50] analyzes video tapes to understand players' cognitive learning when playing a video game. Therefore, for objectives 2 and 3, we assume that different patterns of game play may lead to different learning outcomes.

In portraying children's engagement in the game for objectives 2 and 3, we have adapted the MDA (mechanics, dynamics, aesthetics) framework of Hunicke and colleagues [51]. Aesthetics, which "describes the desirable emotional responses evoked in the player when she interacts with the game system" [51], was not included in our paper as it was unobserved. We define player's reactions to incentives (labeled as incentives) and behaviors in learning about foods (labeled as food learning) as two separate components, as opposed to the paper by Hunicke and colleagues [51], which includes these two in mechanics. Therefore, in this paper, we define game mechanics as the various avatar activities and game features that were experienced by the player (excluding incentives and food learning), game dynamics as the observed responses when the players interact with game mechanics, game strategies as players' reactions to incentives, and food learning as players' choices in learning food-related knowledge in the game.

Methods

Intervention

The intervention used in this study, fooya!, is a mobile gaming app available on iOS and Android platforms that employs mechanisms of implicit learning [33]. Based on experiments and hypotheses derived from pediatric neuropsychology [52], fooya! was developed to deliver therapeutic entertainment that makes healthy behavior change fun for the children. Through prior pilot studies, fooya! has been shown to deliver positive outcomes with respect to food choices, self-reported dietary choices, and healthy eating intentions [34].

Fooya! is designed as an epic action game with an avatar fighting against robots that represent unhealthy/bad foods. Players present themselves as avatars in fooya!, which currently has 80 game levels that progress with increasing difficulty. In each level, the players' main goals are to maintain a good body shape for the avatar and earn enough coins to win the level and

unlock the next level. The body shape changes according to the avatar's food consumption and physical activity (running/jumping) movements, and the changes affect the speed of the avatar. As the avatar consumes more calories, it becomes heavier and its movements become slower. Incentives such as highly prized coins can be earned by destroying the robots (composed of bad foods) using ammo (short for ammunition, composed of bad foods), which implicitly conveys the idea of getting rid of bad foods. Another way to earn the coins, at twice the rate, is by hitting the robots using a shield obtained by consuming good foods (Figure 1a). This feature implicitly conveys the idea that good foods can provide protection from the negative effects of unhealthy foods. After a level is finished, the screen will display a performance summary (Figure 1b) as well as all the food items encountered by the player in that level. The player can choose to click and read these nutrition facts (Figure 1c). Winning a level unlocks the next level, and the player can choose either to replay a level or proceed to the next level (Figure 1d). Throughout the game, the players were not explicitly taught what the healthy and unhealthy foods are and what foods the avatar should consume. The app has no conversation built into the game at any stage, hence explicitly teaching children about healthy/unhealthy foods using language features is not possible. It is designed to be language neutral and has no literacy requirement for use.

Figure 1. Screenshots of Fooya: (a) player is shielded by the bubble after consuming water, (b) performance summary after completing a level, (c) nutrition facts of a chosen food, and (d) level selection.

Experiment and Data

An RCT with pre- and posttreatment measurements was conducted to achieve the objectives of this study. FriendsLearn Inc provided technology access, data collection and extraction, and management support [53]. Power analysis indicated that, with an effect size of 0.50, alpha level of .05, and power of 0.80, at least 102 children needed to be recruited [54,55]. The RCT was conducted in Chennai, India, and had a sample of 104 participants aged 10 to 11 years (same grade) recruited from 3 classes in 3 schools with similar socioeconomic background, one from each school. All students in the participating classes were eligible to join the study. Approvals and availability of class schedules to complete the experiment were confirmed with the school principals and grade teachers prior to the study design. Additionally, all participating students and their parents provided informed consent, and students were randomly assigned to control and treatment groups (Table 1). Demographics (gender, BMI, baseline food preference) were not significantly different between the control and treatment groups (P>.17). Institutional review board approvals from the Center for Media Studies in India, Carnegie Mellon University, and Hofstra University were also obtained. This RCT was retrospectively registered on ClinicalTrials.gov [NCT04082195]. No alterations were made to the experiment design after the RCT began.

Table 1. Sample size, random assignments, and demographic distribution.

Demographics	Treatment group schools (n=52)			Control group schools (n=52)			*P* value[a]
	A (n=20)	B (n=16)	C (n=16)	A (n=19)	B (n=17)	C (n=16)	
Male, n (%)	11 (55)	8 (50)	6 (38)	13 (68)	9 (53)	12 (75)	.63
Female, n (%)	9 (45)	8 (50)	10 (63)	6 (32)	8 (47)	4 (25)	—
BMI (kg/m^2), mean (SD)	17.3 (3.20)	18.1 (4.13)	—[b]	19.2 (6.26)	19.2 (4.42)	—	.17
GoodBase[c], n (%)	8 (50)	3 (23)	3 (21)	12 (86)	5 (29)	7 (50)	.29

[a]*P* values were reported from chi-square tests for gender and GoodBase and 2-sample *t* test for BMI for the pooled sample from schools A and B only, because data from school C were excluded when performing statistical tests for objectives 1 and 3.

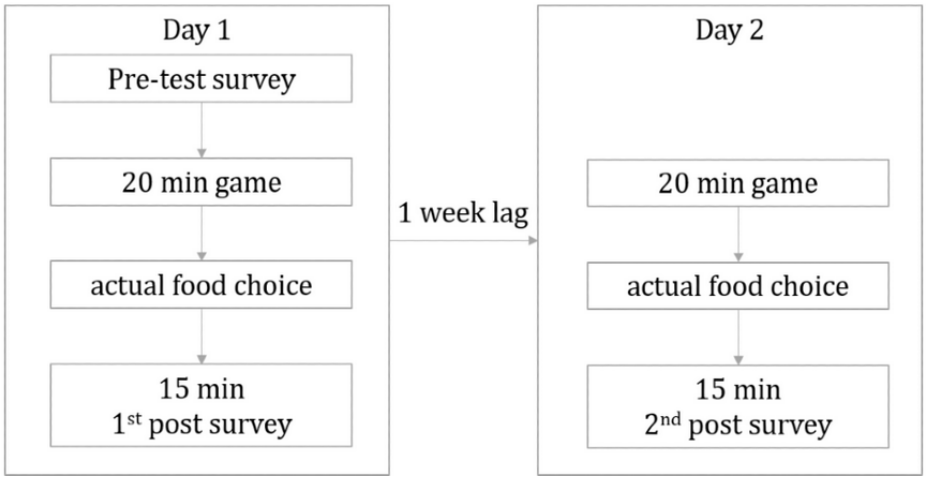
[b]BMI for school C was not collected.

[c]GoodBase captures the number of students who reported good food as their favorite food at baseline. There are some missing values for this variable.

A structured pretest questionnaire was used to collect demographics, food eaten in the previous week, food preferences, and use/frequency of playing video games. Students in the treatment group played fooya! for 20 minutes in a continuous session during the school day, while the control group played a board game (Uno), which does not deliver any knowledge about healthy eating. After the game, the children were shown 3 pairs of food items (healthy and unhealthy) offered from 3 categories—drinks (water and Nimbooz, a carbonated soft drink), savory snacks (cashews and Lays potato chips), and sweet snacks (raisins and a 5 Star chocolate bar)—and asked to choose 2 items. These food items were chosen due to their popularity among children of that age in that region of India. Food selection was conducted in a separate room from the game playing classroom, except in school C, where students made their food choices in the game playing room with other students present. Because this posed the risk of potential contamination of their food choices, school C was excluded from all analyses related to food choices. After the food choice, students were allowed to consume the selected food items.

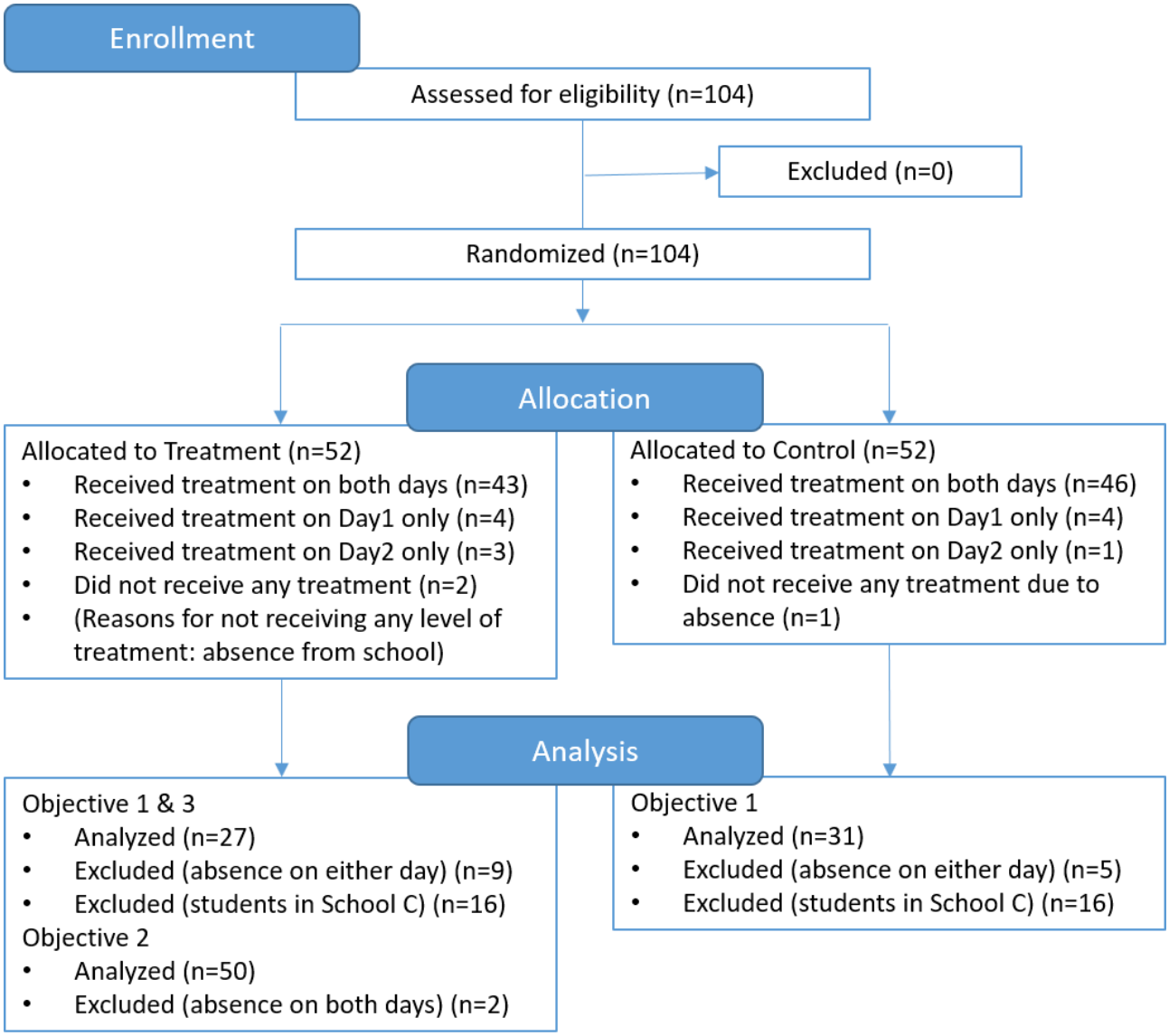
Research staff who recorded the food choice responses were blinded to whether the kids were from the control or treatment group, and participants were blinded to the purpose of the study. Finally, a posttest questionnaire was administered to collect information similar to the pretest questionnaire but with additional information on nutritional knowledge. One week later, the same procedure was administered again with the exception of the pretest survey. Therefore, the children were exposed to the intervention twice, 1 week apart, each for 20 minutes, with children mostly picking up the game from where they left off in the first session. Figure 2 illustrates the study procedures.

Figure 2. Study procedures.



This study is based on a subset of the data collected during the above-mentioned RCT. We analyzed data collected from three sources: survey data (one question each from pretest and posttest), observations of the children's actual food choices, and game telemetry that included about 65,000 game play actions/clicks. Objectives 1 and 3 were investigated using the data collected from schools A and B, and objective 2 was analyzed using the data from all three schools. Survey data that were irrelevant to our objectives were not analyzed or reported. Figure 3 provides the details of the flow of participants and data being analyzed.

Figure 3. Consolidated Standards of Reporting Trials flow diagram of participants through the trial.



Analysis for Objective 1

To analyze the main effect on actual food choice (objective 1), the total number of good foods chosen in two exposures by the treatment and control groups was compared. We also examined the main effect on food knowledge, measured by one survey question asking the students to identify healthy food items from a list of 4 items. The total number of correctly identified healthy food items in two surveys was compared between the control and treatment groups. As secondary outcomes, the effects on the changes of both major outcomes were assessed. Mann-Whitney *U* tests were applied to avoid normality assumption. Cohen *d* was calculated to obtain effect size.

Analysis for Objective 2

To explore children's game play patterns (objective 2), we focused on the telemetry data generated by the participants assigned to the treatment group. Our goal was to understand how each game feature was used and summarize play patterns when children encountered the features for the first time. Therefore, only the data associated with the first encounter of a level by each child was analyzed even if children repeated playing a level a second time.

First of all, each player's telemetry data were processed and encoded according to a feature labeling scheme by which every feature in the game was assigned a distinct feature label. Labels S and E were used to indicate the initiation and termination, respectively, of the level. Labels 1 to 5 represent players' events in the game associated with their avatar (Table 2): (1) consuming good/healthy foods to shield themselves from bad/unhealthy food robots (good food shield), (2) destroying a robot using a shield (destroy by shield), (3) shooting bad food ammo (bad food shot), (4) destroying a robot using bad food shots (destroy by shot), and (5) being hit with bad foods thrown by robots (hit). These labeled events were ordered chronologically to form a sequence based on their recorded timestamps. For example, the time stamped events shown in Table 3 will result in a sequence of S124331E, which has a sequence length of 8 with 6 transitions. More details are described elsewhere [34], and our method of sequence labeling is consistent with existing studies [32,56,57]. These sequences were used to model the process of game playing at any level across all students as a discrete,

time-homogeneous, first-order Markov chain with 7 states, one for each labeled event.

In addition to describing the game play patterns as a model, we further defined game play measures using the MDA framework [51]. The first component in the MDA model is game mechanics, measured as the highest level played in the 2 sessions (Level_max) and average sequence length across levels played (Avg_Seqlen). The second component is game dynamics, which describes how swiftly, or dynamically, the player navigates different features. This is measured by the average number of transitions to a different state in a level across all levels played (Avg_transition). The third component is game strategies, which capture how the player reacts to game rewards. Some might not react to the reward program, others may work hard to get rewards, and still others may figure out the approach to earn rewards strategically. This is measured by the proportion of robots destroyed by shield (Sum2_Sum2+4). To make sure that this is a legal fraction, we use  for each player. This measure is chosen to reflect the fact that destroying a robot by shield is highly rewarded with coins and thus demonstrates how strategically a player is amassing rewards. The last game component is food learning. We measure this by counting the average number of good and bad food facts read per level (AvgGFact and AvgBFact, respectively). Descriptive statistics are calculated to summarize these variables.

Table 2. Definitions of each feature label.

Label	Definition
S	Player starts the level
E	Player finishes the level
1	Player consumes good food to generate shield
2	Robot killed by good food shields
3	Player shoots bad food ammo
4	Robot killed by bad food shots
5	Player hit by bad food robot

Table 3. Example of timestamped data and labeling producing sequence S124331E.

Level	Date	Time	Status	Label/State
Level 1	7/10	09-20-19-775	Player starts the level	S
Level 1	7/10	09-20-26-482	Player consumes good food to generate shield	1
Level 1	7/10	09-20-28-956	Robot killed by good food shields	2
Level 1	7/10	09-21-10-714	Robot killed by bad food shots	4
Level 1	7/10	09-21-10-814	Player shoots bad food ammo	3
Level 1	7/10	09-21-10-894	Player shoots bad food ammo	3
Level 1	7/10	09-21-14-717	Player consumes good food to generate shield	1
Level 1	7/10	09-21-20-281	Player finishes the level	E

Analysis for Objective 3

To examine the effects of player engagement on actual food choices (objective 3), we first identified the game play measures (the 6 measures for the MDA framework defined in objective 2) that were associated with the behavioral outcome. To this end, we conducted Kendall τ correlation tests (to avoid the linearity assumption) and stepwise feature selection using the step() function in the stats package in the R software (R Foundation for Statistical Computing). The identified variables were then included as explanatory variables in regression models for the total number of good foods chosen in 2 sessions (GoodChoice, ranging from 0 to 4) [36]. We considered modeling GoodChoice as normal and Poisson distributions. The normal distribution was chosen as it had the lowest Akaike information criterion for the distribution fit (values for normal and Poisson were 79.64 and 84.23, respectively). Given the fact that GoodChoice was a count measure, a Poisson regression was also run. We used the data pooled from the 2 sessions to avoid spillover effects. Similar to objective 2, only the first-encounter sequences were used in this analysis.

Results

Objective 1

For objective 1, we found strong evidence of the positive main effect of the mobile game on food choices (Table 4). Specifically, children from the treatment group chose 1.38 more good foods during the 2 exposures, on average, than the control group (treatment 2.48, control 1.10; $P<.001$; Cohen $d=1.25$). The power to detect such an effect size, with an alpha level of .05 and sample size of 58, was more than 99% [54,55]. The number of healthy foods correctly identified by the treatment group in the posttest survey question was also significantly

higher than the control group (treatment 7.3, control 6.94; *P*=.048; Cohen *d*=.25).

When comparing the change in both outcomes (day 2 – day 1) between the two groups, no statistically significant results were found, suggesting that the impact of the intervention on healthy food choices after the first session may be sustained after the second session. Additionally, the nonsignificant effects of the game on the change in food choices may have been partly due to the low dosage the children received in the experiment—only 40 minutes in two sessions combined, played 1 week apart. This result supports both the theory and literature, described in the Introduction, that game play experience is associated with health behavioral outcomes. Next, we investigated the potential patterns during game play that were related to this observed positive main effect with analyses of objectives 2 and 3.

Table 4. Treatment effects on food choice and knowledge.

Measures	Actual food choice				Good food ID			
	T[a] (n=27), mean (SD)	C[b] (n=31), mean (SD)	Statistics		T (n=27), mean (SD)	C[a] (n=31), mean (SD)	Statistics	
			P value[c]	*d*[d]			*P* value	*d*
Number of good foods chosen/identified in 2 days (0 to 4 for choice; 0 to 8 for ID; day 1 + day 2)[e]	2.48 (1.19)	1.10 (1.08)	<.001	1.25	7.30 (1.64)	6.94 (1.31)	.048	0.25
Change in the number of good foods chosen/identified (–2 to 2 for choice; –8 to 8 for ID; day 2 – day 1)[e]	–0.11 (–0.93)	–0.13 (1.02)	.92	0.02	–0.26 (0.59)	0.23 (1.12)	.15	–0.54

[a]T: treatment.

[b]C: control.

[c]*P* values are reported from Mann-Whitney tests.

[d]d: Cohen *d*.

[e]Only the responses available on both days for schools A and B were used for comparison.

Objective 2

A total of 842 sequences were collated from the approximately 65,000 game play actions/clicks. Out of these, 115 sequences from repeated plays of random game levels by 38 children were removed, resulting in 727 first-encounter sequences for analysis. The telemetry data analysis indicated high variability in game play patterns among the children (Table 5). For example, children played 15 levels in 2 sessions, on average, ranging from a low of only 2 levels to a high of 23 levels. Coefficient of variation indicated that the largest variance among the variables was in the proportion of scoring activities that were highly rewarded (Sum2_Sum2+4), with an average of 0.17, ranging from 0.003 to 0.98. The descriptive summary in Table 6 shows that there was also a large variation in feature use among children and types of features. The variation was illustrated by the sample sequences shown in Figure 4. While child 15, in Figure 4, did not perhaps know that they could throw bad foods at level 2, they learned to use all features by level 5. In contrast, child 16 was only hit by the robots and could not find a way to fight back. Across different players and their levels of play, the number, type, and sequences of states greatly varied and was reflective of their game experience.

Table 5. Summary statistics of game play measures.

MDA[a] components and game play measures	Min[b]	Mean	Max[c]	Variance	Coefficient of variation
Game mechanics					
Level_max	2.00	14.78	23.00	21.73	31.54
Avg_Seqlen	8.96	65.77	145.42	893.28	45.44
Game dynamics					
Avg_transition	3.64	22.51	40.37	76.70	38.91
Game strategies					
Sum2_Sum2+4	0.003	0.17	0.98	0.05	131.53
Food learning					
AvgGFact	0.00	0.43	1.69	0.18	98.67
AvgBFact	0.00	0.12	0.60	0.02	117.85

[a]MDA: mechanics, dynamics, aesthetics.

[b]Min: minimum.

[c]Max: maximum.

Table 6. Frequency by state for selected levels.

State	Level 1 (n=50)[a]			Level 10 (n=41)			Level 20 (n=8)[b]		
	Min[c]	Max[d]	Mean	Min	Max	Mean	Min	Max	Mean
1	0	5	2.16	0	6	2.70	2	3	2.63
2	0	57	2.02	0	106	2.98	0	19	2.88
3	0	136	32.94	0	175	61.68	0	116	67.25
4	0	17	5.64	0	22	9.15	0	21	8.75
5	0	14	1.76	0	10	4.27	0	11	6.00

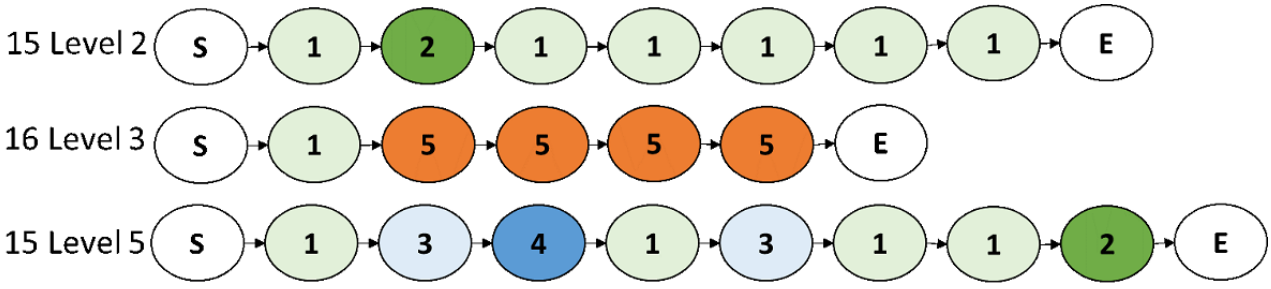
[a]Two children assigned to the treatment group were absent for both exposures.

[b]This is the last level with number of players n>5.

[c]Min: minimum.

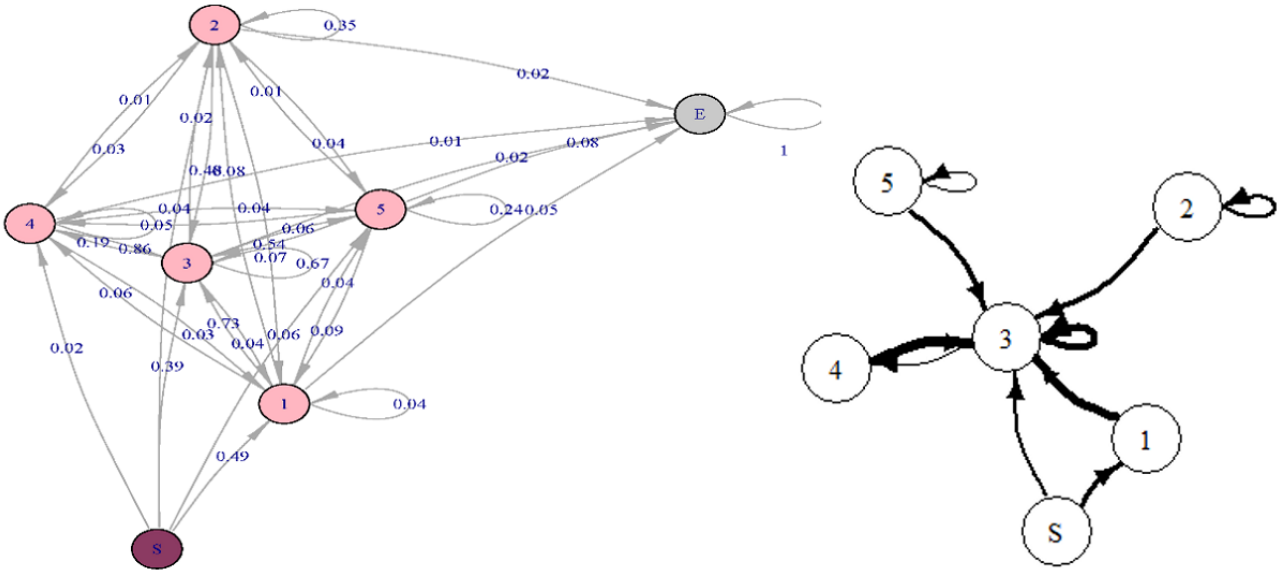
[d]Max: maximum.

Figure 4. Sample sequences.



From the Markov chain model in Figure 5, we observe that overall, bad food shot (state 3) played a central role in the game. It could immediately follow all other activities but was mostly followed only by itself (P_{33}=.67) or robots being destroyed by food shots (P_{34}=.19). We also note that good food shield was seldom followed by destroy by shield (P_{12}=.08), suggesting that children failed to identify and take advantage of this highly rewarded transition throughout the experiment. This finding also confirms the theory that children experience many different learning styles demonstrated by their varied game play patterns.

Figure 5. Markov chain for all children who played level 1 (n=50): (left) detailed plot with all transitions and (right) simplified plot with transition probabilities >0.1. The thickness of the lines is proportional to the transition probability.



Objective 3

Table 7 shows the correlation between the outcome (GoodChoice) and the game play measures. Only AvgBFact had a significant association with GoodChoice (Kendall τ=–.32, *P*=.04). However, both AvgBFact and AvgGFact were retained after the stepwise feature selection procedure. To explore the possible associations, both variables were included in the regression model. The results are shown in models 1 (assumed normal) and 3 (assumed Poisson) in Table 8. The normality assumption for model 1 is satisfied using the Shapiro-Wilk normality test.

Table 7. Kendall rank correlation coefficients and significance test results.

Measures	GoodChoice	*P* value
Level_max[a]	–0.11	0.48
Avg_Seqlen[b]	0.10	0.50
Avg_transition[c]	–0.01	0.95
Sum2_Sum2+4[d]	–0.08	0.58
AvgGFact[e]	0.17	0.25
AvgBFact[f]	–0.32	0.04

[a]Level_max: maximum level played.

[b]Avg_Seqlen: average sequence length across levels played.

[c]Avg_transition: average number of transitions to a different state in a level across all levels played.

[d]Sum2_Sum2+4: proportion of robots destroyed by shield.

[e]AvgGFact: average number of nutrition facts of good foods read in a level.

[f]AvgBFact: average number of nutrition facts of bad foods read in a level.

Table 8. Regression results for GoodChoice[a] (n=25).

Measures	Assumed normal				Assumed Poisson			
	Model 1, coefficient (robust SE)	*P* value	Model 2, coefficient (robust SE)	*P* value	Model 3, coefficient (robust SE)	*P* value	Model 4, coefficient (robust SE)	*P* value
Intercept	2.89 (0.31)	<.001	2.44 (0.39)	<.001	1.06 (0.09)	<.001	0.86 (0.14)	<.001
AvgGFact[b]	1.12 (0.91)	0.23	1.15 (0.62)	0.08	0.44 (0.17)	0.01	0.49 (0.15)	0.001
AvgBFact[c]	–9.51 (3.33)	0.01	–8.38 (3.60)	0.03	–4.25 (1.53)	0.01	–3.81 (1.45)	0.01
GoodBase[d]	—[e]	—	0.92 (0.38)	0.03	—	—	0.36 (0.13)	0.01
Goodness of fit	Adjusted R^2=.31	—	Adjusted R^2=.47	—	AIC[f]: 83.55	—	AIC: 83.59	—

[a]Outcome variable: number of good foods chosen in day1+day2 (0-4). Only the data available on both days in schools A and B were used.

[b]AvgGFact: average number of nutrition facts of good foods read in a level.

[c]AvgBFact: average number of nutrition facts of bad foods read in a level.

[d]GoodBase: binary variable indicating whether the player's baseline preference was healthy or not healthy.

[e]Not applicable.

[f]AIC: Akaike information criterion.

Consistent with Kendall rank test, models 1 and 3 agreed that reading more facts about bad foods is associated with worse food choices. Furthermore, we see that AvGFact may have a positive effect, as shown in model 3. Although the data were collected from a randomized experiment, it is still a concern that these relationships may be due to the children's food preference at baseline. To control for their food preference at baseline, we included a binary variable GoodBase indicating whether the player's baseline preference was healthy (coded as 1) or not healthy (coded as 0). This variable was obtained from an open-ended question in the pretest survey asking the children to write down their favorite food. These open-ended answers were then coded by two raters independently. Disagreements were resolved via a consensus meeting. If no consensus could be reached, a third rater made the final decision. Cohen kappa for overall raters' agreement was 0.72 before consensus and 0.91 after consensus [58].

The results after including the baseline food preference are shown in models 2 (assumed normal) and 4 (assumed Poisson). Not surprisingly, baseline preference had a positive effect on actual food choice. However, the negative effect of AvgBFact still remained, and the potential positive effect of AvgGFact became evident.

Discussion

Principal Findings

Video games with implicit learning strategies, perceived by children as a fun activity and not a learning tool, present a great opportunity to change children's health behaviors by delivering relevant knowledge implicitly. However, without explicit education, whether and how the implicit knowledge can be internalized and result in behavioral change is uncertain. Despite the null effect of some video games found in recent studies [9,12,14], our study provides new evidence by comparing fooya! with a standard board game not designed to deliver implicit knowledge on healthy eating. Our result showed that a mobile game that embeds implicit learning in the game mechanism can positively impact children's actual food choices, consistent with the theory that the tacit knowledge learned from game play can be used to make good decisions [15] and that game play experience can translate to real-world food choices [35-37]. This result is also consistent with empirical studies that have examined the effect of serious video games on health behaviors [21,22].

Furthermore, our analysis of children's play patterns showed significant variations in game play among participants, confirming theories that suggest different children can experience different learning styles [41-43,45]. Sharma and colleagues [7] also found a high variation in time spent in game play and levels played. To the best of our knowledge, our study presents the first attempt to find support for a wide range of learning patterns among players that are associated with different outcomes. Our results show that food choice was not influenced by levels played, same as reported by Sharma and colleagues [7]. Instead, it was influenced by the food facts read in the game: reading more facts about unhealthy foods was associated with more unhealthy food choices and vice versa. This finding is counterintuitive, as one would expect that reading facts about unhealthy foods would lead to decreased choice of unhealthy foods. While our game did not suggest whether the food was good or bad, one possible explanation is that the children were not influenced by the content of the food facts but the impression (the number of views) of the food fact. This is in line with the phenomenon found in studies investigating the effects of advergames [59], in which food choice was affected by the foods advertised in the game. This may be a result of priming [60], which posits that a person's response to the current stimuli (actual food choice) can be influenced by previous stimuli (facts read). Despite the negative effect of bad food facts, our subjects looked up more good food facts than bad food facts, which was one of the drivers of the overall positive effect of the game.

Our results provide several implications for game design. First, serious health games with implicit education strategies as the primary mechanism can be an effective intervention for improving healthy eating behaviors. Players are able to enjoy

the game without noticing that they are being educated. Second, video game designers may want to limit the display of unhealthy foods in the game, especially games that present players with nutritional facts on foods [9], unless some form of protective or explanatory messages are also in place. Furthermore, game designers may also want to insert displays of food facts within the level to make it easier for players to associate the food facts with the change in the avatar's body shape and speed. Third, video games may be coupled with other interventions, whether in-game or not, for children with different game play patterns such as those who read many bad food facts in the game. This personalization of the game or intervention is made possible by associating the actual food choice with game play behaviors measured and quantified using detailed game telemetry.

Limitations and Future Research

This research provides several directions for future extensions by reviewing its limitations. The major limitation of this study is the small sample size. The authors recruited 104 students, but due to the potential bias in the behavioral outcome in one school, data from only 58 students were analyzed for objectives 1 and 3. Due to the observed large effect size, the small sample size is not a problem for objective 1. However, for objective 3, the small sample size prevents us from including the full set of game play measures and controls such as gender and other demographics. It also limits the reliability of regression estimates, especially for the 3-variable models. This small sample size, short game sessions with fixed duration, and the homogeneous subject groups also limit the generalizability of our results. However, these limitations are not unique to this paper, as can be seen in recent game-related literature. For example, existing literature reporting studies with sample size of 46 [61], 20-minute game exposure time [62], or subjects recruited from a single institution [61,63] still provide novel and useful insights. Second, the fact that the actual food choices were made immediately after playing the game limits our observed effect to be the immediate effect. Future longitudinal studies could impose a delay before making the actual food choice to minimize the priming effect and examine whether the game effect is sustained over a longer period and how long it takes for the effect to be internalized.

Third, the kids were not required to consume the chosen food items on site, raising a concern that the strategically thinking kids may choose good foods just to leave a good impression. This concern may be minimal, as the students were not told the purpose of the study. Their choices were also not observed by the people whose opinions they care about, such as their classmates or teacher. With ample sample size and data, future longitudinal studies may further minimize this concern by controlling for students' academic performance, which can be used as a proxy for smartness. Fourth, the choice of a board game (Uno) as a control only allows us to compare the effect of fooya! with a nonserious game. With established effects of games with implicit education strategies, future studies may further examine whether fooya! is preferred to other modes of education on healthy nutrition. Fifth, the BMI in our control group is slightly higher than that in our treatment group, raising a concern that we may have overestimated the size of our main effect. However, the difference in BMI is not statistically significant between the two groups. Last, we use the game play pattern of the first play as a proxy for learning styles.

Future studies may establish the association between learning styles and game play patterns such that tailoring of the intervention can happen before game play. With the overall positive effect found in this study, future studies may also carefully craft and compare the effect of individual implicit education strategies with the results of explicit education components. In addition, analyzing the game play patterns of repeated plays of a level to understand players' established routines rather than first-encounter effects is another avenue for exploration. With more high-quality evidence obtained from longitudinal studies that address these limitations, there is potential for such digital interventions to act as a digital vaccine for pediatric obesity.

Conclusions

Implicit and gamified learning about healthy eating delivered via a mobile app can significantly improve children's food choices immediately after the game. While additional scientific evidence is needed to confirm that such apps can serve as a digital vaccine with long-term impact, this study provides novel insights about the potential drivers of the observed positive short-term effect. We measured and quantified the variation in children's game play patterns that may be associated with children's health behavior outcomes. Specifically, the positive main effect may be strengthened when players read more nutrition facts about healthy foods in the game. However, the effect is compromised when the players read more facts about unhealthy foods. Results of this study provide promising directions for the use of a viable alternative to improve children's eating habits and help address the pediatric overweight and obesity epidemic in the long term. Future research with large scale, longitudinal RCTs is needed to fully understand how game play patterns with their underlying learning styles may facilitate desired behavioral outcomes.

Acknowledgments

The authors thank Satya Gopal, PN Ramya and Niveditha Srinivasan from Ethiraj College, Chennai, and Nikitha Carmel, Chanchal Yadav, and Rena Singhi for their help with the RCT; and Palak Narang, Sross Gupta, Disha Gupta, and Ketan Deshmukh at Carnegie Mellon University for their assistance in data processing and preliminary analysis. The authors greatly appreciate the comments and suggestions from the anonymous reviewers of JMIR, which have significantly improved the manuscript. This study was partly funded by the Remala Foundation to support graduate students assisting with data processing and preliminary analysis.

Authors' Contributions

RP and YCKL defined the study objectives, conducted the data analysis, and wrote and edited the manuscript. UBK and PK designed and conducted the RCT and edited the manuscript. BSP provided the mobile game intervention and technical support and edited the manuscript. BSP, VV, and SA helped to conduct the trial and collect the data. NS provided clinical knowledge and contributed to the design of the RCT.

Conflicts of Interest

Coauthor BSP is the founder and CEO of FriendsLearn, the company that produced the mobile game used in this study, but provided no financial support and did not influence the design or results of this study. Other authors declared there are no competing interests.

Editorial Notice

This randomized study was only retrospectively registered due to the lack of awareness about the need to register it by the co-authors who conducted the RCT in Chennai. They obtained the required approvals from the Institutional Review Board in India, but did not recognize the need to register the trial prior to recruiting participants. They affirm that no alterations to the experiment design were made after the RCT was initiated. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V. 1. 6. 1).
[PDF File (Adobe PDF File), 1324 KB - mhealth_v8i11e15717_app1.pdf]

References

1. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011-2012. JAMA 2014 Feb 26;311(8):806-814. [doi: 10.1001/jama.2014.732] [Medline: 24570244]
2. Wright DR, Prosser LA. The impact of overweight and obesity on pediatric medical expenditures. Appl Health Econ Health Policy 2014 Apr;12(2):139-150. [doi: 10.1007/s40258-014-0088-7] [Medline: 24652198]
3. DeSmet A, Van Ryckeghem D, Compernolle S, Baranowski T, Thompson D, Crombez G, et al. A meta-analysis of serious digital games for healthy lifestyle promotion. Prev Med 2014 Dec;69:95-107. [doi: 10.1016/j.ypmed.2014.08.026] [Medline: 25172024]
4. Baranowski T, Baranowski J, Thompson D, Buday R, Jago R, Griffith MJ, et al. Video game play, child diet, and physical activity behavior change a randomized clinical trial. Am J Prev Med 2011 Jan;40(1):33-38 [FREE Full text] [doi: 10.1016/j.amepre.2010.09.029] [Medline: 21146765]
5. Baranowski T, Baranowski J, Cullen KW, Marsh T, Islam N, Zakeri I, et al. Squire's Quest! Dietary outcome evaluation of a multimedia game. Am J Prev Med 2003 Jan;24(1):52-61. [Medline: 12554024]
6. DeSmet A, Liu Y, De Bourdeaudhuij I, Baranowski T, Thompson D. The effectiveness of asking behaviors among 9-11 year-old children in increasing home availability and children's intake of fruit and vegetables: results from the Squire's Quest II self-regulation game intervention. Int J Behav Nutr Phys Act 2017 Apr 21;14(1):51 [FREE Full text] [doi: 10.1186/s12966-017-0506-y] [Medline: 28431547]
7. Sharma SV, Shegog R, Chow J, Finley C, Pomeroy M, Smith C, et al. Effects of the Quest to Lava Mountain computer game on dietary and physical activity behaviors of elementary school children: a pilot group-randomized controlled trial. J Acad Nutr Diet 2015 Aug;115(8):1260-1271. [doi: 10.1016/j.jand.2015.02.022] [Medline: 25912520]
8. Putnam MM, Richmond EM, Brunick KL, Wright CA, Calvert SL. Influence of a character-based app on children's learning of nutritional information: should apps be served with a side of media characters? Games Health J 2018 Jan 24;7(2):121-126. [doi: 10.1089/g4h.2017.0116] [Medline: 29364706]
9. Hermans RCJ, van den Broek N, Nederkoorn C, Otten R, Ruiter ELM, Johnson-Glenberg MC. Feed the Alien! The effects of a nutrition instruction game on children's nutritional knowledge and food intake. Games Health J 2018 Jun;7(3):164-174. [doi: 10.1089/g4h.2017.0055] [Medline: 29634366]
10. Schneider KL, Ferrara J, Lance B, Karetas A, Druker S, Panza E, et al. Acceptability of an online health videogame to improve diet and physical activity in elementary school students: “Fitter Critters”. Games Health J 2012 Aug;1(4):262-268 [FREE Full text] [doi: 10.1089/g4h.2012.0009] [Medline: 24761317]
11. Majumdar D, Koch PA, Lee H, Contento IR, Islas-Ramos ADL, Fu D. “Creature-101”: a serious game to promote energy balance-related behaviors among middle school adolescents. Games Health J 2013 Oct;2(5):280-290 [FREE Full text] [doi: 10.1089/g4h.2013.0045] [Medline: 24761326]

12. Bell BM, Martinez L, Gotsis M, Lane HC, Davis JN, Antunez-Castillo L, et al. Virtual Sprouts: a virtual gardening pilot intervention increases self-efficacy to cook and eat fruits and vegetables in minority youth. Games Health J 2018 Apr;7(2):127-135 [FREE Full text] [doi: 10.1089/g4h.2017.0102] [Medline: 29394102]
13. Holzmann SL, Schäfer H, Groh G, Plecher DA, Klinker G, Schauberger G, et al. Short-term effects of the serious game. Nutrients 2019 Aug 30;11(9):2031 [FREE Full text] [doi: 10.3390/nu11092031] [Medline: 31480257]
14. Folkvord F, Hage G, Theben A. Can Garfield make children eat healthier: a cluster-randomized controlled trial to examine the effect of a serious health game on children's eating behavior (preprint). JMIR Serious Games 2019. [doi: 10.2196/preprints.16888]
15. Reber AS. Implicit learning and tacit knowledge. J Experimental Psychol 1989;118(3):219-235. [doi: 10.1037/0096-3445.118.3.219]
16. Lumosity. URL: https://www.lumosity.com [accessed 2020-10-19]
17. Hardy JL, Nelson RA, Thomason ME, Sternberg DA, Katovich K, Farzin F, et al. Enhancing cognitive abilities with comprehensive training: a large, online, randomized, active-controlled trial. PLoS ONE 2015 Sep 2;10(9):e0134467. [doi: 10.1371/journal.pone.0134467]
18. Kpolovie PJ. Effect of lumosity training schedule on recall of nonsense syllables. Adv Soc Sci Res J 2018;5(10):206-234 [FREE Full text] [doi: 10.14738/assrj.510.5290]
19. Ciavarro C, Dobson M, Goodman D. Implicit learning as a design strategy for learning games: alert hockey. Comput Human Behav 2008;24(6):2862-2872 [FREE Full text] [doi: 10.1016/j.chb.2008.04.011]
20. Rowe E, Asbell-Clarke J, Baker R, Eagle M, Hicks A. Assessing implicit science learning in digital games. Comput Human Behav 2017;76:617-630 [FREE Full text] [doi: 10.1016/j.chb.2017.03.043]
21. Charlier N, Zupancic N, Fieuws S, Denhaerynck K, Zaman B, Moons P. Serious games for improving knowledge and self-management in young people with chronic conditions: a systematic review and meta-analysis. J Am Med Inform Assoc 2016 Jan;23(1):230-239. [doi: 10.1093/jamia/ocv100] [Medline: 26186934]
22. Chow CY, Riantiningtyas RR, Kanstrup MB, Papavasileiou M, Liem GD, Olsen A. Can games change children's eating behaviour? A review of gamification and serious games. Food Qual Prefer 2020 Mar;80:103823. [doi: 10.1016/j.foodqual.2019.103823]
23. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. JMIR Serious Games 2014;2(2):e9 [FREE Full text] [doi: 10.2196/games.3413] [Medline: 25654660]
24. Baranowski T, Baranowski J, Thompson D, Buday R. Behavioral science in video games for children's diet and physical activity change: key research needs. J Diabetes Sci Technol 2011 Mar 01;5(2):229-233 [FREE Full text] [doi: 10.1177/193229681100500204] [Medline: 21527086]
25. Thompson D. Designing serious video games for health behavior change: current status and future directions. J Diabetes Sci Technol 2012 Jul 01;6(4):807-811 [FREE Full text] [Medline: 22920806]
26. Spook J, Paulussen T, Kok G, van Empelen P. Evaluation of a serious self-regulation game intervention for overweight-related behaviors ("Balance It"): a pilot study. J Med Internet Res 2016 Sep 26;18(9):e225 [FREE Full text] [doi: 10.2196/jmir.4964] [Medline: 27670222]
27. Gagné A, El-Nasr M, Shaw C. A deeper look at the use of telemetry for analysis of player behavior in rts games. 2011 Presented at: 2011 International Conference on Entertainment Computing (ICEC 2011); 2011; Vancouver. [doi: 10.1007/978-3-642-24500-8_26]
28. Moura D, el-Nasr M, Shaw C. Visualizing and understanding players' behavior in video games: discovering patterns and supporting aggregation and comparison. 2011 Presented at: 2011 ACM SIGGRAPH symposium on video games; 2011; Vancouver. [doi: 10.1145/2018556.2018559]
29. Owen V, Ramirez D, Salmon A, Halverson R. Capturing learner trajectories in educational games through ADAGE (assessment data aggregator for game environments): a click-stream data framework for assessment of learning in play. 2014 Presented at: 2014 American Educational Research Association Annual Meeting; 2014; Philadelphia.
30. Westera W, Nadolski R, Hummel H. Serious gaming analytics: what students' log files tell us about gaming and learning. Int J Serious Games 2014 Jun 19;1(2). [doi: 10.17083/ijsg.v1i2.9]
31. Smith S, Blackmore K, Nesbitt K. A meta-analysis of data collection in serious games research. In: A Meta-Analysis of Data Collection in Serious Games Research: Methodologies for Performance Measurement, Assessment, and Improvement. Berlin: Springer; 2015:31-55.
32. Loh CS, Sheng Y. Measuring the (dis-)similarity between expert and novice behaviors as serious games analytics. Educ Inf Technol 2013 Jun 10;20(1):5-19. [doi: 10.1007/s10639-013-9263-y]
33. Fooya. URL: https://fooya.com/ [accessed 2020-10-19]
34. Padman R, Gupta D, Sri Prakash B, Krishnan C, Panchatcharam K. An exploratory analysis of game telemetry from a pediatric mHealth intervention. Stud Health Technol Inform 2017;245:183-187. [Medline: 29295078]
35. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Washington: American Psychological Association; 1986.
36. Garris R, Ahlers R, Driskell JE. Games, motivation, and learning: a research and practice model. Simul Gaming 2002 Dec 01;33(4):441-467. [doi: 10.1177/1046878102238607]

37. Gupta S, Bostrom R. Research note: an investigation of the appropriation of technology-mediated training methods incorporating enactive and collaborative learning. Inform Syst Res 2013 Jun;24(2):454-469. [doi: 10.1287/isre.1120.0433]
38. Schank R, Cleary C. Engines for Education. Hoboken: Lawrence Erlbaum Associates, Inc; 1995.
39. Brown AL, Kane MJ. Preschool children can learn to transfer: learning to learn and learning from example. Cogn Psychol 1988 Oct;20(4):493-523. [doi: 10.1016/0010-0285(88)90014-x] [Medline: 3191668]
40. Folkvord F, Anschütz DJ, Buijzen M, Valkenburg PM. The effect of playing advergames that promote energy-dense snacks or fruit on actual food intake among children. Am J Clin Nutr 2013 Feb;97(2):239-245. [doi: 10.3945/ajcn.112.047126] [Medline: 23269821]
41. McDermott PA, Beitman BS. Standardization of a scale for the study of children's learning styles: structure, stability, and criterion validity. Psychol Schs 1984 Jan;21(1):5-14. [doi: 10.1002/1520-6807(198401)21:1<5::aid-pits2310210102>3.0.co;2-b]
42. Hassan MA, Habiba U, Majeed F, Shoaib M. Adaptive gamification in e-learning based on students' learning styles. Interact Learn Envir 2019 Mar 17:1-21. [doi: 10.1080/10494820.2019.1588745]
43. Wouters P, van der Meulen E. The role of learning styles in game-based learning. Int J Game Based Learn 2020;10(1):54-69. [doi: 10.4018/ijgbl.2020010104]
44. Coffield F, Moseley D, Hall E, Ecclestone K, Coffield F. Learning Styles and Pedagogy in Post-16 Learning: A Systematic and Critical Review. London: Learning and Skills Research Centre; 2004.
45. Robinson JE, Gray JL. Cognitive style as a variable in school learning. J Educ Psychol 1974;66(5):793-799. [doi: 10.1037/h0037487]
46. Hwang G, Hsu T, Hsieh Y. Impacts of different smartphone caption/subtitle mechanisms on English listening performance and perceptions of students with different learning styles. Int J Hum Comput Interact 2018 Nov 08;35(4-5):333-344. [doi: 10.1080/10447318.2018.1543091]
47. Buckley P, Doyle E. Individualising gamification: an investigation of the impact of learning styles and personality traits on the efficacy of gamification using a prediction market. Comput Educ 2017 Mar;106:43-55. [doi: 10.1016/j.compedu.2016.11.009]
48. Saxe GB. Studying cognitive development in sociocultural context: the development of a practice‐based approach. Mind Cult Act 1994 Jun;1(3):135-157. [doi: 10.1080/10749039409524666]
49. Saxe G. Culture and Cognitive Development: Studies in Mathematical Understanding. Philadelphia: Taylor & Francis; 2015.
50. Nasir NS. Individual cognitive structuring and the sociocultural context: strategy shifts in the game of dominoes. J Learn Sci 2005 Jan;14(1):5-34. [doi: 10.1207/s15327809jls1401_2]
51. Hunicke R, Leblanc M, Zubek R. A formal approach to game design and game research. 2004 Presented at: Proceedings of AAAI Workshop on Challenges in Game AI; 2004; San Jose.
52. Bruce AS, Bruce JM, Black WR, Lepping RJ, Henry JM, Cherry JBC, et al. Branding and a child's brain: an fMRI study of neural responses to logos. Soc Cogn Affect Neurosci 2012 Sep 20;9(1):118-122. [doi: 10.1093/scan/nss109]
53. FriendsLearn. URL: https://friendslearn.com/ [accessed 2020-10-19]
54. Chow S, Shao J, Wang H, Lokhnygina Y. Sample Size Calculations in Clinical Research. Boca Raton: CRC Press; 2017.
55. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Philadelphia: Taylor & Francis; 2013.
56. Zheng K, Padman R, Johnson MP, Diamond HS. An interface-driven analysis of user interactions with an electronic health records system. J Am Med Inform Assoc 2009;16(2):228-237 [FREE Full text] [doi: 10.1197/jamia.M2852] [Medline: 19074301]
57. Lin Y, Padman R. Process visibility analysis in ambulatory care: a simulation study with rfid data. Stud Health Technol Informat 2013;192:768-772.
58. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 2016 Jul 02;20(1):37-46. [doi: 10.1177/001316446002000104]
59. Dias M, Agante L. Can advergames boost children's healthier eating habits? A comparison between healthy and non-healthy food. J Consumer Behav 2011 Jun 28;10(3):152-160. [doi: 10.1002/cb.359]
60. Molden DC. Understanding priming effects in social psychology: what is "social priming" and how does it occur? Soc Cogn 2014 Jun;32 Supplement:1-11. [doi: 10.1521/soco.2014.32.supp.1]
61. Robert P, Manera V, Derreumaux A, Ferrandez Y Montesino M, Leone E, Fabre R, et al. Efficacy of a web app for cognitive training (MeMo) regarding cognitive and behavioral performance in people with neurocognitive disorders: randomized controlled trial. J Med Internet Res 2020 Mar 11;22(3):e17167. [doi: 10.2196/17167]
62. Shum AK, Lai ES, Leung WG, Cheng MN, Wong HK, So SW, et al. A digital game and school-based intervention for students in Hong Kong: quasi-experimental design. J Med Internet Res 2019 Apr 05;21(4):e12003 [FREE Full text] [doi: 10.2196/12003] [Medline: 30950795]
63. Phungoen P, Promto S, Chanthawatthanarak S, Maneepong S, Apiratwarakul K, Kotruchin P, et al. Precourse preparation using a serious smartphone game on advanced life support knowledge and skills: randomized controlled trial. J Med Internet Res 2020 Mar 09;22(3):e16987 [FREE Full text] [doi: 10.2196/16987] [Medline: 32149711]

Abbreviations

Avg_Seqlen: average sequence length across levels played
Avg_transition: average number of transitions to a different state in a level across all levels played
AvgBFact: average number of bad food facts read per level
AvgGFact: average number of good food facts read per level
GoodBase: binary variable indicating whether the player's baseline preference was healthy or not healthy
GoodChoice: total number of good foods chosen in the 2 sessions
Level_max: highest level played in the 2 sessions
MDA: mechanics, dynamics, aesthetics
RCT: randomized controlled trial
Sum2_Sum2+4: proportion of robots destroyed by shield

Edited by G Eysenbach; submitted 01.08.19; peer-reviewed by T Ueno, P Giabbanelli, T Baranowski; comments to author 07.02.20; revised version received 24.06.20; accepted 03.10.20; published 18.11.20.

Please cite as:
Kato-Lin YC, Kumar UB, Sri Prakash B, Prakash B, Varadan V, Agnihotri S, Subramanyam N, Krishnatray P, Padman R
Impact of Pediatric Mobile Game Play on Healthy Eating Behavior: Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(11):e15717
URL: http://mhealth.jmir.org/2020/11/e15717/
doi:10.2196/15717
PMID:33206054

Original Paper

Behavior Change Techniques in Wrist-Worn Wearables to Promote Physical Activity: Content Analysis

Peter Düking[1], MSc; Marie Tafler[2], BSc; Birgit Wallmann-Sperlich[3], PhD; Billy Sperlich[1], PhD; Sonja Kleih[2], PhD

[1]Integrative and Experimental Exercise Science, Department of Sport Science, Würzburg, Germany

[2]Department of Psychology I, University of Würzburg, Würzburg, Germany

[3]Institute for Sport Sciences, University of Würzburg, Würzburg, Germany

Corresponding Author:
Peter Düking, MSc
Integrative and Experimental Exercise Science
Department of Sport Science
University of Würzburg
Würzburg, 97082
Germany
Phone: 49 931 31 84792
Email: peterdueking@gmx.de

Abstract

Background: Decreasing levels of physical activity (PA) increase the incidences of noncommunicable diseases, obesity, and mortality. To counteract these developments, interventions aiming to increase PA are urgently needed. Mobile health (mHealth) solutions such as wearable sensors (wearables) may assist with an improvement in PA.

Objective: The aim of this study is to examine which behavior change techniques (BCTs) are incorporated in currently available commercial high-end wearables that target users' PA behavior.

Methods: The BCTs incorporated in 5 different high-end wearables (Apple Watch Series 3, Garmin Vívoactive 3, Fitbit Versa, Xiaomi Amazfit Stratos 2, and Polar M600) were assessed by 2 researchers using the BCT Taxonomy version 1 (BCTTv1). Effectiveness of the incorporated BCTs in promoting PA behavior was assessed by a content analysis of the existing literature.

Results: The most common BCTs were goal setting (behavior), action planning, review behavior goal(s), discrepancy between current behavior and goal, feedback on behavior, self-monitoring of behavior, and biofeedback. Fitbit Versa, Garmin Vívoactive 3, Apple Watch Series 3, Polar M600, and Xiaomi Amazfit Stratos 2 incorporated 17, 16, 12, 11, and 11 BCTs, respectively, which are proven to effectively promote PA.

Conclusions: Wearables employ different numbers and combinations of BCTs, which might impact their effectiveness in improving PA. To promote PA by employing wearables, we encourage researchers to develop a taxonomy specifically designed to assess BCTs incorporated in wearables. We also encourage manufacturers to customize BCTs based on the targeted populations.

(JMIR Mhealth Uhealth 2020;8(11):e20820) doi:10.2196/20820

KEYWORDS

cardiorespiratory fitness; innovation; smartwatch; technology; wearable; eHealth; mHealth

Introduction

Various forms of physical activity (PA) reduce the incidence of noncommunicable diseases, obesity, and mortality. However, according to the World Health Organization (WHO), levels of physical inactivity are increasing with approximately 28% of adults failing to meet PA guidelines [1]. Therefore, measures to increase PA are urgently needed.

Behavioral PA interventions (eg, employing cognitive and behavioral techniques to modify and increase PA behavior) successfully increase PA [2]. However, these interventions target smaller groups of mostly previously motivated participants. The eHealth and mobile health (mHealth) [3] solutions employing wearable sensors (wearables) may encourage various populations to increase their levels of PA on a larger scale. Wearables monitor certain components of PA via surrogate markers (eg, movement acceleration converted to metabolic equivalents or electrical signals converted to heart rate) [4] and provide biofeedback [5], thereby potentially assisting in elevating PA. A recent review concludes that

wearables have the potential to increase PA participation as long as wearables are the primary component of an intervention or part of a broader intervention [6]. Additionally, the WHO aims to endorse digital health concepts [7]. In Germany, physicians are permitted to prescribe digital health solutions if proven to be effective [8]. At the same time, wearable-assisted interventions may be more cost-effective than traditional interventions [9].

To promote PA, different behavior change techniques (BCT) can be incorporated in wearables, which likely have different outcomes for promoting PA [3]. The selection of wearables using appropriate BCTs based on specific research questions and goals of healthy behavior is particularly crucial in the continuously growing wearable market. Manufacturers are releasing new models with rapidly changing features at least once every year, creating more choices for researchers and consumers.

Currently, little is known about how wearables differ from each other and which technologies are more effective in increasing levels of PA. Therefore, this study aimed to examine which BCTs targeting PA behavior are incorporated in the commercially available high-end wearables.

Methods

Wearables

For our analysis, we chose wearables manufactured by leading companies in the market [10]. The 5 wrist-worn wearables were the Apple Watch Series 3 (Apple Inc), Fitbit Versa (Fitbit Inc), Garmin Vívoactive 3 (Garmin), Polar M600 (Polar Electro Oy), and Xiaomi Amazfit Stratos 2 (Huami Technology). Each wearable was installed as instructed by the respective manufacturer and was synchronized with the companion app for smartphones (Apple iPhones).

Coding Procedure

Wearables and companion apps were coded using the BCT Taxonomy Version 1 (BCTTv1), which was previously employed in similar studies [11-13]. The BCTTv1 is explained in detail by Michie et al [14]. Briefly, the BCTTv1 incorporated 93 nonredundant techniques, grouped into 16 hierarchical clusters in total, each coded using a dichotomous score of either 0 or 1, indicating nonpresence or presence, respectively [14].

Each wearable was worn by 2 researchers (MT and PD) for 1 week. The 2 researchers were well acquainted with the handling of the wearables, using the companion apps, and employing the BCTTv1. The researchers completed the training on the BCTTv1 website [15] before the analysis. Interrater reliability assessing each wearable incorporating a BCT was calculated using a kappa statistic in SPSS 22.0 (IBM Corp). Magnitude of agreement was interpreted as per the following criteria: 0.00=poor, 0.01-0.20=slight, 0.21-0.40=fair, 0.41-0.60=moderate, 0.61-0.80=substantial, and 0.81-1.00=almost perfect [16].

Coding disagreements were resolved by a discussion between the researchers. In case of disagreement, a third researcher's opinion (SK) was included to resolve the disagreement.

In line with the aim of this study, BCTs targeting PA were examined, while the feedback on other factors (eg, sleep or diet) was ignored. The researchers were instructed to include periods of physical inactivity as well as those of PA into the assessment week to verify the corresponding feedback by wearables.

Evaluating the Effectiveness of BCTs in the Wearables to Promote PA

To evaluate the potential effects of incorporated BCTs, we employed a previous list [9] created for the same purpose. The list is based on a meta-analysis [17,18], meta-regression [19], and systematic reviews [20-22] as well as recommendations from the US Preventive Services Taskforce [23]. These BCTs were marked with checkmarks in Table 1. As in the earlier study [9], we used this list to count the number of effective BCTs, which were incorporated in each wearable to promote PA.

Results

Table 1 summarizes the different BCTs incorporated in the 5 different wearables. Techniques from the taxonomy not immanent in any of the systems were excluded from the table. One disagreement between the original 2 researchers was solved by the opinion of the third one. The interrater reliability was almost perfect (Cohen kappa=0.965).

Table 1. Behavior change techniques incorporated in different wrist-worn wearables.

BCTs[a]	Proven effectiveness to promote physical activity	Apple Watch Series 3	Fitbit Versa	Garmin Vívoactive 3	Polar M600	Xiaomi Amazfit Stratos 2	Incorporations, N
Goal setting (behavior) (item 1.1)	✓	✓	✓	✓	✓	✓	5
Barrier identification/problem solving (item 1.2)	✓						0
Action planning (item 1.4)	✓	✓	✓	✓	✓	✓	5
Review behavior goal(s) (item 1.5)	✓	✓	✓	✓	✓	✓	5
Discrepancy between current behavior and goal (item 1.6)		✓	✓	✓	✓	✓	5
Commitment (item 1.9)	✓						0
Feedback on behavior (item 2.2)	✓	✓	✓	✓	✓	✓	5
Self-monitoring of behavior (item 2.3)	✓	✓	✓	✓	✓	✓	5
Biofeedback (item 2.6)	✓	✓	✓	✓	✓	✓	5
Social support (unspecified) (item 3.1)	✓	✓	✓	✓	✓	✓	5
Social support (emotional) (item 3.3)			✓	✓	✓		3
Instruction on how to perform the behavior (item 4.1)	✓		✓[b]	✓			1 (2)
Information about health consequences (item 5.1)	✓		✓	✓	✓	✓	4
Information about social and environmental consequences (item 5.3)	✓						0
Monitoring of emotional consequences (item 5.4)					✓		1
Information about emotional consequences (item 5.6)	✓						0
Demonstration of the behavior (item 6.1)	✓		✓[b]	✓			1 (2)
Social comparison (item 6.2)	✓	✓	✓	✓	✓	✓	5
Prompts/cues (item 7.1)	✓	✓	✓	✓	✓	✓	5
Behavioral practice/rehearsal (item 8.1)	✓		✓[b]	✓			1 (2)
Graded tasks (item 8.7)	✓	✓	✓	✓			2
Credible source (item 9.1)			✓	✓		✓	3
Nonspecific reward (item 10.3)	✓	✓	✓	✓	✓		4
Social reward (item 10.4)		✓	✓	✓	✓	✓	5
Nonspecific incentive (item 10.6)		✓	✓	✓			3

BCTs[a]	Proven effectiveness to promote physical activity	Apple Watch Series 3	Fitbit Versa	Garmin Vívoactive 3	Polar M600	Xiaomi Amazfit Stratos 2	Incorporations, N
Self-reward (item 10.9)	✓						—[c]
Adding objects to the environment (item 12.5)		✓	✓	✓	✓	✓	5
Reward approximation (item 14.4)	✓		✓				1
Situation-specific reward (item 14.6)		✓	✓	✓			3
Verbal persuasion about capability (item 15.1)		✓		✓			2
Focus on past success (item 15.3)	✓	✓	✓	✓		✓	4
Self-talk (item 15.4)	✓						—
BCT clusters, n		10	13	13	8	10	—
Incorporated BCTs, n		18	24	24	16	15	—
Incorporated BCTs with proven effectiveness [17-23], n		12	17	16	11	11	—

[a]BCT: behavior change technique.

[b]Available as a paid add-on feature.

[c]Not applicable.

Incorporated BCTs

Out of the 93 BCTs analyzed by the BCTTv1, 26 different BCTs were incorporated in the 5 wearables. On average, 19 BCTs (range 15-24) were incorporated in the wearables. Fitbit Versa and Garmin Vívoactive 3 incorporated the most BCTs (n=24), followed by Apple Watch Series 3 (n=18), Polar M600 (n=16), and Xiaomi Amazfit Stratos 2 (n=15). Due to technical issues with the Xiaomi Amazfit Stratos 2 device, we could not evaluate the BCT item of social support (emotional) (item 3.3); thereby, that item was not marked as incorporated with a checkmark in Table 1. Multimedia Appendix 1 provides detailed information about how often a BCT was incorporated in each wearable.

According to Lyons et al [9], 23 BCTs are effective in promoting PA. Out of these BCTs, Garmin Vívoactive 3 (n=16) and Fitbit Versa (n=14 + 3 paid BCTs) incorporate most BCTs, followed by Apple Watch Series 3 (n=12), Polar M600 (n=11), and Xiaomi Amazfit Stratos 2 (n=11). Only Fitbit Versa and Garmin Vívoactive 3 included the BCT items of instruction on how to perform behavior and behavioral practice/rehearsal. None of the wearables included the following BCT items: barrier identification/problem solving, self-reward, self-talk, commitment, information about social and environmental consequences, and information about emotional consequences.

Clusters

On average, 11 (range 8-13) clusters were incorporated in the 5 wrist-worn wearables. Fitbit Versa and Garmin Vívoactive 3 incorporated most clusters (n=13), followed by Apple Watch Series 3 (n=10), Amazfit Stratos 2 (n=10), and Polar M600 (n=8). The 3 most common clusters were goals and planning (cluster 1) (n=4), feedback and monitoring (cluster 2) (n=3), and reward and threat (cluster 10) (n=3).

Discussion

This study was designed to examine which BCTs were incorporated in the leading high-end wearables to promote PA. Our major findings were as follows:

1. The most common BCTs were goal setting (behavior), action planning, review behavior goal(s), discrepancy between current behavior and goal, feedback on behavior, self-monitoring of behavior, biofeedback, social support (unspecified), social comparison, prompts/cues, social reward, and adding objects to the environment.
2. Wearables often incorporate the same BCTs according to the BCTTv1. However, Fitbit Versa and Garmin Vívoactive 3 provided the most and Xiaomi Amazfit Stratos 2 provided the least number of BCTs.
3. Fitbit Versa (n=17) and Garmin Vívoactive 3 (n=16) offered the most BCTs, which showed to be effective to promote PA, while Xiaomi Amazfit Stratos 2 and Polar M600 had the least number of BCTs (n=11).

The number of incorporated BCTs in this study is in line with previous research examining incorporated BCTs within earlier versions of the wearables tested herein [11-13]. In studies comparing different wearables, Fitbit incorporated a higher number of BCTs than those incorporated by the Garmin [12] or Polar [11,13] wearables.

Absolute numbers of BCTs incorporated in wearables might not effectively inform about which wearables seem suitable to increase levels of PA. People's decision making can be

deteriorated by information overload and may even result in negatively perceived stress [24,25]. Consequently, everyone requires an optimal level of information quantity. In this regard, fewer but more effective BCTs may be preferable to promote PA [9].

A combination of BCTs (with proven effectiveness) may maximize effectiveness in promoting PA. A meta-analysis revealed that (1) "provide information about behavior–health link" combined with "prompt intention formation" (mean effect size, g=0.46) and (2) "provide information about behavior–health link" combined with "provide information on consequences" and "use of follow-up prompts" (mean effect size, g=0.44) were the most successful BCTs to alter health behavior [26]. When converted to the BCTTv1 taxonomy, the equivalent BCTs are (1) information about health consequences (item 5.1) combined with action planning (item 1.4), and (2) information about health consequences (item 5.1) with prompts/cues (item 7.1), respectively. A total of 4 devices tested herein (those manufactured by Fitbit, Garmin, Polar, and Xiaomi) incorporated all these BCTs; while the device manufactured by Apple incorporated all the BCTs except the information about health consequences (item 5.1).

Although the aforementioned BCTs and their combinations are effective in promoting PA on group level, effective BCTs may differ for specific populations and on an individual level. In fact, individual characteristics (eg, age, PA level, and personality) play a key role in a person's choices and continuous device usages [27]. Consequently, it may be worthwhile to customize the incorporated BCTs in wearables to meet individual characteristics for PA promotion.

On an individual level, the effectiveness of BCTs and the choice of a wearable depend on practical factors. For example, BCTs incorporated in wearables are only effective when worn, which (among other factors) depend largely on personal preferences, including design and texture, battery life, and handling [27]. Additionally, sensor data that prompt different BCTs need reliable information [28]. For example, if a wearable uses inertial measurement units and optical heart rate sensors to calculate a person's energy expenditure (a surrogate marker of PA) for prompting different BCTs, the sensor data and algorithms need to be as accurate as possible. Otherwise, the BCTs may mislead and ultimately fail personal PA goals. In our experience, companies often do not disclose the information of sensor data and algorithms applied to prompt a certain BCT. Transparency in this regard would assist in understanding the prompting of BCT and identifying preferred wearables to promote PA.

An effective increase of PA also depends on the user's awareness of PA and its impact on health [29]. Using wearables without such knowledge might not represent an optimal option to improve PA. Although a recent meta-analysis showed that wearables improve PA (ie, measured by the number of daily steps, moderate to vigorous PA, and energy expenditure) discretely, a combination of wearables with other BCTs (eg, telephone counseling or group-based education) shows better results in increasing levels of PA [6].

Since wearables are commercially available, they can be used by consumers without consulting an expert opinion (eg, opinion of an exercise physiologist or a health expert). Potentially, the lack of professional guidance can induce unfavorable or dangerous PA participation (eg, by increasing PA too rapidly).

Limitations

Although the taxonomy we applied to assess BCTs in wearables was employed previously [9], this taxonomy was not directly developed for machine to person (ie, wearable to consumer) interaction but for person to person (ie, psychologist to patient) interaction. This modification results in some difficulties in the application of this taxonomy. For example, different surrogate markers of PA (eg, number of steps or energy expenditure) can be employed to prompt the same BCT. To the best of our knowledge, it remains unclear which surrogate marker is most suitable as a basis for BCTs to improve PA.

Additionally, the BCTTv1 taxonomy does not allow certain functions, for example, evaluating how a BCT is visualized, evaluating how frequently BCTs are implemented (eg, daily report and hourly feedback), and determining whether a BCT is interactive (eg, a BCT disappears from the screen when an increase in PA is detected). However, we assume that these aspects affect the effectiveness of the BCTs and hence, that of the respective wearables in altering PA behavior among users.

Since the number of wearables promoting PA is rapidly increasing, we urge researchers to develop a taxonomy covering these aspects to assess the preferred BCTs employed by wearables.

Some BCTs displayed by the wearables addressed sedentary behavior as well as PA (ie, prompts/cues). Sedentary behavior is recognized as an independent behavior of PA [30,31]. In this study, we did not separate BCTs in terms of PA behavior and sedentary behavior. However, sedentary behavior and PA follow unique intervention logic [31,32], and research suggests that different factors influence each behavior. Future research should analyze the incorporated BCTs in wearables for sedentary behavior, separately of PA. Targeting both behaviors can have greater health benefits.

The shelf life of consumer-grade wearables, such as that of the devices we analyzed in this study, is short. Manufacturers frequently release new models, while older ones disappear from the market. Consequently, frequent assessment of new wearables, ideally before the market release (as recommended elsewhere [28]) is warranted.

Conclusions

Based on our analysis, we conclude that out of all the tested wearables, Fitbit Versa (n=14+3 paid) and Garmin Vívoactive 3 (n=16) included most of the BCTs with proved effectiveness, followed by Apple Watch 5 (n=12), Polar M600 (n=11), and Xiaomi Amazfit Stratos 2 (n=11). Out of all the tested wearables, Garmin Vívoactive 3 and Fitbit Versa might be the most promising wearables to promote PA from a psychological perspective. However, future studies need to evaluate the effectiveness of these devices in experimental studies along with the validity and reliability of variables obtained by these devices. Since no specific taxonomy is available to investigate BCTs incorporated in digital health tools, including wearables

for promoting PA, we advise developing such a taxonomy given the global urgency to improve PA and the increasing popularity and availability of wearables. Since the effectiveness of BCTs is affected by individual characteristics, we recommend manufacturers to allow customization of BCTs.

Acknowledgments

This publication was supported by the Open Access Publication Fund of the University of Wuerzburg.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of the behavior change techniques incorporated in the wearables tested herein.
[PDF File (Adobe PDF File), 39627 KB - mhealth_v8i11e20820_app1.pdf]

References

1. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Glob Health 2018 Oct;6(10):e1077-e1086 [FREE Full text] [doi: 10.1016/S2214-109X(18)30357-7] [Medline: 30193830]
2. Conn VS, Hafdahl AR, Mehr DR. Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. Am J Public Health 2011 Apr;101(4):751-758 [FREE Full text] [doi: 10.2105/AJPH.2010.194381] [Medline: 21330590]
3. Hosseinpour M, Terlutter R. Your Personal Motivator is with You: A Systematic Review of Mobile Phone Applications Aiming at Increasing Physical Activity. Sports Med 2019 Sep;49(9):1425-1447 [FREE Full text] [doi: 10.1007/s40279-019-01128-3] [Medline: 31144235]
4. Düking P, Hotho A, Holmberg H, Fuss FK, Sperlich B. Comparison of Non-Invasive Individual Monitoring of the Training and Health of Athletes with Commercially Available Wearable Technologies. Front Physiol 2016;7:71 [FREE Full text] [doi: 10.3389/fphys.2016.00071] [Medline: 27014077]
5. Düking P, Holmberg H, Sperlich B. Instant Biofeedback Provided by Wearable Sensor Technology Can Help to Optimize Exercise and Prevent Injury and Overuse. Front Physiol 2017;8:167 [FREE Full text] [doi: 10.3389/fphys.2017.00167] [Medline: 28420998]
6. Brickwood K, Watson G, O'Brien J, Williams AD. Consumer-Based Wearable Activity Trackers Increase Physical Activity Participation: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2019 Apr 12;7(4):e11819 [FREE Full text] [doi: 10.2196/11819] [Medline: 30977740]
7. Be Healthy, Be Mobile: A Handbook on How to Implement mHypertension. Geneva: World Health Organization and International Telecommunication Union; 2020.
8. Deutscher Bundestag. Entwurf eines Gesetzes für eine bessere Versorgung durch Digitalisierung und Innovation. Deutscher Bundestag. 2019. URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/D/DVG_Bundestag.pdf [accessed 2019-11-29]
9. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014 Aug 15;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
10. IDC. Worldwide Wearables Shipments Surge 94.6% in 3Q 2019 Led by Expanding Hearables Market, Says IDC. IDC. 2019 Dec 09. URL: https://www.idc.com/getdoc.jsp?containerId=prUS45712619 [accessed 2020-03-09]
11. Chia GLC, Anderson A, McLean LA. Behavior Change Techniques Incorporated in Fitness Trackers: Content Analysis. JMIR Mhealth Uhealth 2019 Jul 23;7(7):e12768 [FREE Full text] [doi: 10.2196/12768] [Medline: 31339101]
12. Duncan M, Murawski B, Short CE, Rebar AL, Schoeppe S, Alley S, et al. Activity Trackers Implement Different Behavior Change Techniques for Activity, Sleep, and Sedentary Behaviors. Interact J Med Res 2017 Aug 14;6(2):e13 [FREE Full text] [doi: 10.2196/ijmr.6685] [Medline: 28807889]
13. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis. JMIR Mhealth Uhealth 2016 Apr 27;4(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4461] [Medline: 27122452]
14. 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]
15. BCT Taxonomy Online Training. URL: https://www.bct-taxonomy.com/ [accessed 2020-11-02]

16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
17. Williams SL, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? Health Educ Res 2011 Apr;26(2):308-322. [doi: 10.1093/her/cyr005] [Medline: 21321008]
18. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2013 Mar 03;10:29 [FREE Full text] [doi: 10.1186/1479-5868-10-29] [Medline: 23452345]
19. 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]
20. Bird EL, Baker G, Mutrie N, Ogilvie D, Sahlqvist S, Powell J. Behavior change techniques used to promote walking and cycling: a systematic review. Health Psychol 2013 Aug;32(8):829-838 [FREE Full text] [doi: 10.1037/a0032078] [Medline: 23477577]
21. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011 Feb 18;11:119 [FREE Full text] [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
22. Bélanger-Gravel A, Godin G, Vézina-Im LA, Amireault S, Poirier P. The effect of theory-based interventions on physical activity participation among overweight/obese individuals: a systematic review. Obes Rev 2011 Jun;12(6):430-439. [doi: 10.1111/j.1467-789X.2010.00729.x] [Medline: 20331511]
23. Task Force on Community Preventive Services. Recommendations to increase physical activity in communities. Am J Prev Med 2002 May;22(4 Suppl):67-72. [doi: 10.1016/s0749-3797(02)00433-6] [Medline: 11985935]
24. Baker DA. Four Ironies of Self-quantification: Wearable Technologies and the Quantified Self. Sci Eng Ethics 2020 Jun;26(3):1477-1498. [doi: 10.1007/s11948-020-00181-w] [Medline: 31970596]
25. Eppler MJ, Mengis J. The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. The Information Society 2004 Nov;20(5):325-344. [doi: 10.1080/01972240490507974]
26. Dusseldorp E, van GL, van BS, Verheijden MW, van EP. Combinations of techniques that effectively change health behavior: evidence from Meta-CART analysis. Health Psychol 2014 Dec;33(12):1530-1540. [doi: 10.1037/hea0000018] [Medline: 24274802]
27. 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]
28. Sperlich B, Holmberg H. Wearable, yes, but able…?: it is time for evidence-based marketing claims!. Br J Sports Med 2017 Aug;51(16):1240. [doi: 10.1136/bjsports-2016-097295] [Medline: 27986762]
29. Emanuel EJ, Wachter RM. Artificial Intelligence in Health Care: Will the Value Match the Hype? JAMA 2019 Jun 18;321(23):2281-2282. [doi: 10.1001/jama.2019.4914] [Medline: 31107500]
30. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. Int J Behav Nutr Phys Act 2017 Jun 10;14(1):75 [FREE Full text] [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
31. Biddle S. Fit or sit? Is there a psychology of sedentary behaviour? 2011;7(2):5-10.
32. Biddle S, Mutrie N, Gorely T, Blamey A. Interventions for physical activitysedentary behaviour. In: Roberts GC, Treasure D. editors. Advances in Motivation in Sport and Exercise. Champaign, IL: Human Kinetics; 2012:357-386.

Abbreviations

BCT: behavior change technique
BCTTv1: BCT Taxonomy version 1
mHealth: mobile health
PA: physical activity
WHO: World Health Organization

Edited by G Eysenbach; submitted 29.05.20; peer-reviewed by A Henriksen, A Anderson, MI Cajita, D Arigo; comments to author 12.07.20; revised version received 13.08.20; accepted 02.09.20; published 19.11.20.

Please cite as:
Düking P, Tafler M, Wallmann-Sperlich B, Sperlich B, Kleih S
Behavior Change Techniques in Wrist-Worn Wearables to Promote Physical Activity: Content Analysis
JMIR Mhealth Uhealth 2020;8(11):e20820
URL: http://mhealth.jmir.org/2020/11/e20820/
doi:10.2196/20820
PMID:33211023

Original Paper

Pre-Conception Interventions for Subfertile Couples Undergoing Assisted Reproductive Technology Treatment: Modeling Analysis

Régine Steegers-Theunissen[1], MD, PhD; Annemieke Hoek[2], MD, PhD; Henk Groen[3], MD, PhD; Annelies Bos[4], MD, PhD; Grada van den Dool[5], MD; Marieke Schoonenberg[6], MD; Jesper Smeenk[7], MD, PhD; Eva Creutzberg[8], PhD; Loes Vecht[9], MD; Luc Starmans[9], PhD; Joop Laven[1], MD, PhD

[1]Department of Obstetrics and Gynaecology, Erasmus MC, University Medical Center, Rotterdam, Netherlands
[2]Department of Obstetrics and Gynaecology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands
[3]Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands
[4]Department of Reproductive Medicine and Gynecology, University Medical Centre Utrecht, Utrecht, Netherlands
[5]Department of Obstetrics and Gynaecology, Albert Schweitzer Hospital Zwijndrecht, Zwijndrecht, Netherlands
[6]Nij Geertgen Centre for fertility Elsendorp, Elsendorp, Netherlands
[7]Department of Obstetrics and Gynaecology, Elisabeth-TweeSteden Hospital Tilburg, Tilburg, Netherlands
[8]Department of Gynaecology, Ferring BV Hoofddorp, Amsterdam, Netherlands
[9]KPMG, Amsterdam, Netherlands

Corresponding Author:
Régine Steegers-Theunissen, MD, PhD
Department of Obstetrics and Gynaecology
Erasmus MC
University Medical Center
Office EE-2271a, Dr.Molewaterplein 40, 3015 GD
Rotterdam
Netherlands
Phone: 31 107038255
Email: r.steegers@erasmusmc.nl

Abstract

Background: Approximately 1 in 7 couples experience subfertility, many of whom have lifestyles that negatively affect fertility, such as poor nutrition, low physical activity, obesity, smoking, or alcohol consumption. Reducing lifestyle risk factors prior to pregnancy or assisted reproductive technology treatment contributes to the improvement of reproductive health, but cost-implications are unknown.

Objective: The goal of this study was to evaluate reproductive, maternal pregnancy, and birth outcomes, as well as the costs of pre-conception lifestyle intervention programs in subfertile couples and obese women undergoing assisted reproductive technology.

Methods: Using a hypothetical model based on quantitative parameters from published literature and expert opinion, we evaluated the following lifestyle intervention programs: (1) Smarter Pregnancy, an online tool; (2) LIFEstyle, which provides outpatient support for obese women; (3) concurrent use of both Smarter Pregnancy and LIFEstyle for obese women; (4) smoking cessation in men; and (5) a mindfulness mental health support program using group therapy sessions. The model population was based on data from the Netherlands.

Results: All model-based analyses of the lifestyle interventions showed a reduction in the number of in vitro fertilization, intracytoplasmic sperm injection, or intrauterine insemination treatments required to achieve pregnancy and successful birth for couples in the Netherlands. Smarter Pregnancy was modeled to have the largest increase in spontaneous pregnancy rate (13.0%) and the largest absolute reduction in potential assisted reproductive technology treatments. Among obese subfertile women, LIFEstyle was modeled to show a reduction in the occurrence of gestational diabetes, maternal hypertensive pregnancy complications, and preterm births by 4.4%, 3.8%, and 3.0%, respectively, per couple. Modeled cost savings per couple per year were €41 (US $48.66), €360 (US $427.23), €513 (US $608.80), €586 (US $695.43), and €1163 (US $1380.18) for smoking cessation, mindfulness, Smarter Pregnancy, combined Smarter Pregnancy AND LIFEstyle, and LIFEstyle interventions, respectively.

Conclusions: Although we modeled the potential impact on reproductive outcomes and costs of fertility treatment rather than collecting real-world data, our model suggests that of the lifestyle interventions for encouraging healthier behaviors, all are likely to be cost effective and appear to have positive effects on reproductive, maternal pregnancy, and birth outcomes. Further real-world data are required to determine the cost-effectiveness of pre-conception lifestyle interventions, including mobile apps and web-based tools that help improve lifestyle, and their effects on reproductive health. We believe that further implementation of the lifestyle app Smarter Pregnancy designed for subfertile couples seeking assistance to become pregnant is likely to be cost-effective and would allow reproductive health outcomes to be collected.

(JMIR Mhealth Uhealth 2020;8(11):e19570) doi:10.2196/19570

KEYWORDS

fertility, periconception, pregnancy chance, Smarter Pregnancy, cost-effectiveness, nutrition, obesity, IVF treatment; mobile and web-based lifestyle apps

Introduction

Many couples who undergo fertility treatment have multiple lifestyle risk factors that reduce their chances of becoming pregnant [1]. Female lifestyle risk factors such as smoking, alcohol use, and poor diet are inversely associated with fecundity and time to pregnancy, and the effects of these factors increase with increasing BMI [2]. Chronic stress or high anxiety levels reduce fecundability, and high depression scores are also associated with subfertility [3]. Furthermore, evidence is accumulating that poor nutrition, stress, drug use, infection, or exposure to environmental chemicals during prenatal development has a life-long impact on offspring health as shown by the Developmental Origins of Health and Disease paradigm [4]. Obese pregnant women are more likely to have gestational diabetes, hypertensive complications, premature delivery, higher risk of cesarean delivery, and fetal death [5]. Obese men are more likely than men of normal weight to have lower sperm quality [6].

Despite the evidence available, weight-loss interventions in overweight or obese couples prior to fertility treatment remain controversial, as there is limited evidence that these interventions increase the chance of a live birth or reduce pregnancy complications [7,8]. Two well-conducted randomized controlled trials [9,10], however, show that spontaneous conceptions significantly increased with pre-conception weight-loss lifestyle interventions in obese subfertile women and led to lower number of fertility treatments, though neither trial showed an increase in assisted reproductive technology treatment-dependent pregnancies.

Health care budgets are increasingly under pressure due to noncommunicable diseases associated with an aging population, scarcity in the workforce, and rising costs of novel medical technologies. It is, therefore, key that lifestyle interventions aimed at subfertile couples improve reproductive outcomes and in a cost-effective manner.

The aim of this evaluation was to estimate and model the impact of pre-conception lifestyle interventions on the likelihood of the occurrence of pregnancy, and maternal pregnancy and birth outcomes after fertility treatment in subfertile couples and in subfertile obese women, as well as to assess whether these lifestyle interventions are cost-effective.

Methods

Overview

We created a hypothetical model to estimate the effectiveness of pre-conception lifestyle interventions on reproductive, pregnancy, and birth outcomes and potential cost savings in subfertile couples, including a subgroup of subfertile obese women, undergoing in vitro fertilization (IVF), intracytoplasmic sperm injection (ICSI), or intrauterine insemination (IUI) treatment.

Lifestyle Factors Relevant to Fertility Care

Modifiable factors most relevant to fertility treatment, reproductive health, and pregnancy complications with compelling quantifiable clinical published evidence were women's BMI (a marker of health and lifestyle), diet (nutrition intake), physical activity, smoking cessation, and stress [2,9,11-23].

Pre-conception Lifestyle Interventions

Five lifestyle interventions matching the selected lifestyle factors were selected by the fertility experts for inclusion in the cost-effectiveness model (Table 1). These programs included support for changing lifestyle factors deemed most relevant to pre-conception care that were outlined above. The programs used in our model were Smarter Pregnancy [12,20], and LIFEstyle [9,11]. We also modeled smoking cessation in men by using analysis of IVF and ICSI outcomes in men who smoke and men who do not smoke [24], and mindfulness-based mental health support for couples undergoing IVF [16].

Table 1. Lifestyle interventions selected for cost-effectiveness modeling.

Name	Model target population	Description
Smarter Pregnancy intervention [12,20]	Women aged 25-44 years trying to conceive, comprising subfertile couples seeking medical assistance to conceive (IVF[a]/ICSI[b]) and fertile couples	Provides 26 weeks of individual online coaching and information via smartphone tailored to improve nutrition and lifestyle during the pre-conception and pregnancy period in order to improve the health of the reproductive population and subsequent generations
LIFEstyle intervention [9,11]	Obese subfertile women with BMI >29 kg/m^2 seeking fertility treatment	Provides 6 outpatient visits (each 30 minutes long) and 4 telephone consultations (15 minutes) during a 24-week period to provide motivation and support for nutrition (energy restriction of approximately 500 kCal/day) and exercise strategies (10,000 steps per day, 2-3 moderate vigorous exercise sessions per week for weight loss in obese women (BMI >29 kg/m^2) seeking fertility treatment
Combined Smarter Pregnancy and LIFEstyle intervention [9,11,12,20]	Combination of the 2 target audiences mentioned above	Providing both Smarter Pregnancy and LIFEstyle support for obese women (the remaining couples were modeled to receive the Smarter Pregnancy intervention only)
Smoking cessation [24]	Subfertile men who smoke	Comparison of IVF and ICSI outcomes in male smokers and non-smokers from couples seeks reproductive assistance
Mindfulness support [16]	Subfertile women undergoing their first IVF or ICSI cycle	Comparison of IVF outcomes in couples either receiving or not receiving group sessions to teach stress reduction through a mindfulness-based intervention while undergoing IVF treatment

[a]IVF: in vitro fertilization.

[b]ICSI: intracytoplasmic sperm injection.

Model Population and Treatment Policies

The model population comprised couples living in the Netherlands seeking fertility treatment, representing a general subfertile population including a subgroup of obese women. The prevalence of subfertility in men and women living in the Netherlands aged 25-45 years was 0.7% and 2.2%, respectively, representing approximately 15,000 men and 46,000 women who are subfertile (Multimedia Appendix 1). Based on available published data [25,26], we calculated that there were 5400 subfertile obese women, 3200 smoking men, and 13,700 women undergoing their first IVF or ICSI cycle. Based on clinical experience, we assumed approximately 24% of the couples living in the Netherlands seeking assisted reproductive technology treatment received IVF treatment (mean 1.5 cycles), 16% received ICSI (mean 1.5 cycles) in total per couple, and 60% received IUI (mean 3.0 cycles). We did not include couples seeking ovulation induction in our model.

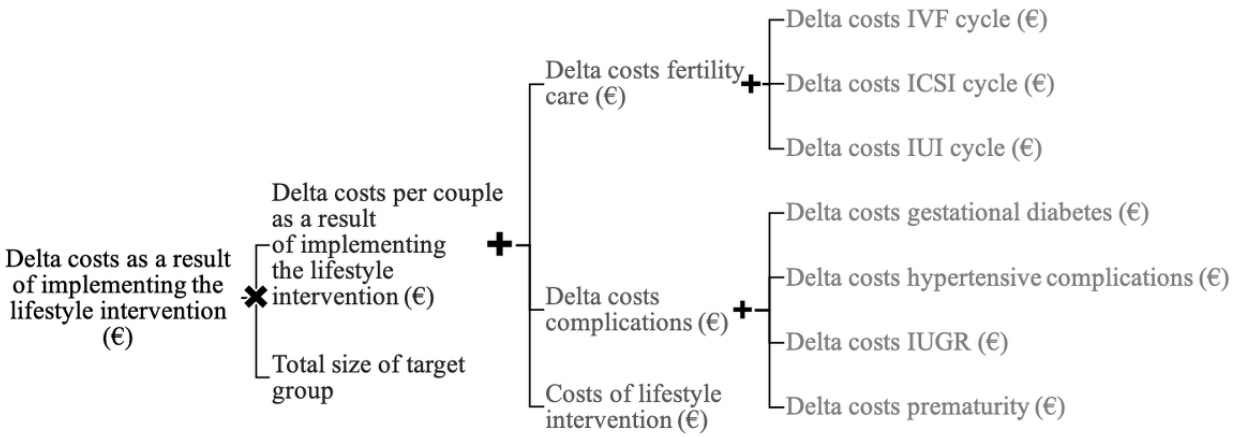
Modeling Clinical Outcomes

For each pre-conception lifestyle intervention program, the published data were reviewed to determine input parameters for the model to estimate the intervention's impact on chance of a spontaneous ongoing pregnancy (a viable pregnancy at week 12), number of IVF, ICSI, and IUI treatments, as well as on pregnancy complications— gestational diabetes, gestational hypertensive complications, and preterm delivery—expressed per couple per year. For each lifestyle intervention, the model only included the fertility treatments and known complications from the lifestyle intervention according to high quality evidence.

Modeling Potential Cost Savings

The general structure of the cost-effectiveness model is depicted in Figure 1. The cost-effectiveness evaluation was performed from a health care perspective and included direct medical costs of the lifestyle intervention, fertility treatments, medication, and any resulting pregnancy. The model (business case) was designed to estimate the cost impact of the selected lifestyle interventions on fertility and obstetric care per subfertile couple, expressed as the difference in costs per patient or per couple and the overall cost difference per year in the Netherlands.

Figure 1. General structure of the cost-effectiveness model for each lifestyle intervention. IUGR: intrauterine growth restriction; IUI: intrauterine insemination; IVF: in vitro fertilization; ICSI: intracytoplasmic sperm injection.



Model Input Parameters

The effects of the respective interventions on each of the selected lifestyle factors were modeled according to available published literature or expert opinion agreed by consensus. Costs for ovulation induction prior to IUI were not included as part of the cost-effectiveness model. Where published data existed, previously estimated costs of pregnancy complications, such as fetal growth restriction, gestational diabetes, gestational hypertensive complications (including preeclampsia), and premature birth, were used.

The general costs for assisted reproductive technology and pregnancy are presented in Multimedia Appendix 1. Costs, indexed to 2016 price levels, for pregnancy, birth and admission until 6 weeks postpartum in singleton and multiple pregnancies conceived after IVF were estimated.

Specific parameters related to each pre-conception lifestyle intervention are presented in Multimedia Appendix 2 including costs of the intervention, the estimated costs of pregnancy outcomes, chances of spontaneous pregnancy, intrauterine growth restriction, and gestational diabetes or hypertension. For each lifestyle intervention that was modeled, the estimated impact on fertility treatment or pregnancy outcomes were only included if there were data to support effects associated with that specific lifestyle intervention.

Results

Modeled Impact of Lifestyle Interventions on Clinical Outcomes

The modeled impact of lifestyle interventions on fertility outcomes, including the reductions in assisted reproduction and pregnancy complications, are presented in Tables 2 and 3. Each of the 5 modeled lifestyle interventions showed a reduction in the number of IVF and ICSI treatments per couple required to achieve a successful ongoing pregnancy. Smarter Pregnancy in the subfertile couples and LIFEstyle in subfertile obese women resulted in 10,800 (23.4%) and 1100 (20.0%) fewer IUI cycles, respectively, among the modeled population for the Netherlands than if no lifestyle intervention was used. Overall, Smarter Pregnancy was modeled to have both the largest increase in the number of spontaneous pregnancies (6000 additional spontaneous pregnancies among the modeled population, which is a 13% reduction in the number of couples who require assisted reproductive technology treatment) and the largest reduction in potential number of assisted reproductive technology treatments per couple (4.7% for in vitro fertilization, 3.1% for intracytoplasmic sperm injection, and 23.4% for intrauterine insemination). Smarter Pregnancy was associated with a reduction of fetal growth restriction by 2.6%, and LIFEstyle was associated with a reduction of gestational diabetes (4.4%), gestational hypertensive complications (3.8%), and preterm delivery (3.0%). There were no published data regarding the effects of smoking cessation or mindfulness on the likelihood of achieving spontaneous pregnancies, or on fetal growth restriction, gestational diabetes, hypertensive, or premature birth pregnancy complication.

Table 2. The effect of each lifestyle intervention on assisted reproductive technology clinical outcomes modeled for the Netherlands.

Intervention	Modeled changes in clinical results (per year)			
	Spontaneous pregnancies, n (%)[a]	IVF[b] treatments, n (%)	ICSI[c] treatments, n (%)	IUI[d] treatments, n (%)
Smarter Pregnancy (SlimmerZwanger)	+6000 (13.0)	–2200 (–4.7)	–1400 (–3.1)	–10,800 (–23.4)
LIFEstyle	+500 (9.9)	–600 (–11.3)	–1000 (–18.9)	–1100 (–20.0)
Smarter Pregnancy + LIFEstyle[e]	+6000 (13.0)	–2500 (–5.5)	–2300 (–5.0)	–10,800 (–23.4)
Smoking cessation (men only)	__[f]	–300 (–0.8)	–100 (–0.4)	—
Mental health/ mindfulness	—	–1600 (–11.8)	–100 (–0.9)	—

[a]Percentage has been calculated as the proportional change in the number of events (spontaneous pregnancy or assisted reproductive technology treatment resulting in pregnancy) for the model population per intervention.

[b]IVF: in vitro fertilization.

[c]ICSI: intracytoplasmic sperm injection.

[d]IUI: intrauterine insemination.

[e]Only for the subset of women with BMI over 29 kg/m^2.

[f]No published data found in literature searches.

Table 3. Effect of each lifestyle intervention on clinical outcomes of pregnancy complications modeled for the Netherlands (modeled per year).

Intervention	Modeled changes in clinical results (per year)			
	IUGR[a], n (% change)[b]	Gestational diabetes, n (% change)	Hypertensive complications, n (% change)	Decreased number of premature births, n (% change)
Smarter Pregnancy (*SlimmerZwanger*)	–1200 (–2.6)	__[c]	—	—
LIFEstyle	—	–200 (–4.4)	–200 (–3.8)	–200 (–3.0)
Smarter Pregnancy + LIFEstyle	–1200 (–2.6)[d]	–200 (–4.4)[d]	–200 (–3.8)[d]	–200 (–3.0)[d]
Smoking cessation (men only)	—	—	—	—
Mental health/ mindfulness	—	—	—	—

[a]IUGR: intrauterine growth restriction.

[b]Percentage has been calculated as the proportional change in the number of events (pregnancy complications) for the model population per intervention).

[c]No published data found in literature searches.

[d]Only for the subset of women with BMI over 29 kg/m^2.

Estimated Cost-Savings of Lifestyle Interventions and Sensitivity Analysis

A summary of the estimated cost savings for each lifestyle intervention is presented in Table 4. The lifestyle intervention that was modeled to have the highest cost saving was LIFEstyle, with an estimated saving of €1163 (US $1380.18 at the time of publication) per couple; however, this intervention is specifically for subfertile obese women. The combination of Smarter Pregnancy and LIFEstyle, for which obese women would have access to both Smarter Pregnancy and LIFEstyle and all other subfertile couples would use Smarter Pregnancy only, was modeled to save €586 (US $695.43) per couple. Across the entire potential target group in the Netherlands, the greatest financial saving would be achieved with the combination program Smarter Pregnancy and LIFEstyle (€27 million, approximately US $32 million), followed by Smarter Pregnancy (€24 million, approximately US $28.4 million) alone. The lifestyle intervention of smoking cessation for men represented the lowest cost saving €41 (US $48.66) per couple in the Netherlands.

Table 4. Estimated cost savings per year of lifestyle interventions (business cases) on assisted reproductive technologies and reduction of pregnancy complications modeled for the Netherlands (an exchange rate of approximately €1=US $1.19 is applicable at the time of publication).

Intervention	Estimated cost benefit per couple[a] (least favorable, most favorable scenario)	Total target group (prevalence)	Total annual target group (incidence)	Estimated overall cost saving (least favorable, most favorable scenario) per year[a]	Estimated total assisted reproductive technology cost saving per year (least favorable, most favorable scenario)[b]
Smarter Pregnancy (*SlimmerZwanger*)	€513 (€100, €2200)	46,000	1191	€24 M (€4.6 M, €101.2 M)	€6 M (€1.2 M, €26.2 M)
LIFEstyle	€1163 (€900, €1600)	5400	1391	€6 M (€4.9 M, €8.6 M)	€1.6 M (€1.3 M, €101 M)
Smarter Pregnancy + LIFEstyle	€586 (€100, €2200)	46,000	1191	€27 M (€4.6 M, €101 M)	€7 M (€4.6 M, €26.4 M)
Smoking cessation (men only)	€41 (€–20, €170)	3200	826	€0.130 M (€–0.064 M, €0.54 M)	€0.034 M (€–0.017 M, €0.140 M)
Mental health/ Mindfulness	€36 (€–190, €500)	13,700	3551	€4.9 M (€–2.6 M, €6.9 M)	€1.3 M (€–0.7 M, €1.8 M)

[a]Overall cost savings include all medical intervention costs for complications.

[b]Includes cost savings for assisted reproductive technology procedures that would no longer be necessary.

Discussion

Principal Results

Using a model population based on subfertile couples and subfertile obese women living in the Netherlands undergoing IVF, ICSI, or IUI, we developed a hypothesis-based model using quantitative parameters from published literature and expert opinion to explore how 5 selected lifestyle intervention programs, each using different approaches and targets, can potentially improve the chances of spontaneous pregnancy, reduce the number of cycles of IVF, ICSI, or IUI treatments, and reduce the chance of adverse maternal pregnancy and birth outcomes. As part of the analysis our model also assessed potential cost savings of the selected lifestyle intervention programs for the model population. In order to estimate the cost savings, we calculated the estimated cost savings per couple for each of the selected lifestyle intervention programs. Multiplying this number with the size of the target group of the selected lifestyle intervention program yielded the total estimated cost savings. Cost savings per couple were modeled for each of the lifestyle programs taking into account the cost savings achieved by reducing the volume of fertility care, as well as cost savings achieved due to a reduction of pregnancy complications. Estimated costs of the lifestyle intervention were subsequently subtracted from these savings, in order to achieve the total cost savings of the lifestyle intervention. For each lifestyle intervention, the model only included the fertility treatments and complications that the lifestyle intervention was known to have an effect upon using available evidence.

Modifiable lifestyle factors were selected based on their positive associations with the likelihood of spontaneous conception or successful fertility treatment and thus would be most relevant for evaluating in terms of potential cost saving. Age was considered to be important but not modifiable at the moment of the health care visit. In men, moderate alcohol consumption, caffeine intake, and scrotal temperature were considered less relevant lifestyle factors and, therefore, were not included in our model. Thus, the most important modifiable factors were maternal BMI (used as a surrogate marker of health and lifestyle), diet, physical activity, smoking, and stress. These factors had sufficient published quantitative data regarding their impact on reproductive health, fertility treatment, and pregnancy complications to develop our model and have also used by others [9,11,12,16,20,24].

In our model, both Smarter Pregnancy and LIFEstyle increased the number of spontaneous pregnancies by 13.0% and 9.9% per couple, respectively, compared with no lifestyle intervention. This supports data from another study that showed that a 1-point increase in a pre-conception dietary risk score was associated with 65% increased chance of ongoing pregnancy [27] and a “Mediterranean”-style diet is likely to improve IVF and ICSI treatment success with an increased probability of pregnancy (odds ratio 1.4, 95% CI 1.0-1.9) [28]. Similarly, in our model, these 2 lifestyle interventions decreased the number of IVF, ICSI, and IUI cycles required for a successful pregnancy. Although there were no data available for the effect of smoking cessation in men and mindfulness mental health support on spontaneous pregnancies or number of IUI treatments, our model suggests that both may decrease the number of IVF and ICSI treatments required for a successful pregnancy.

Our model also showed that validated lifestyle interventions may contribute to a reduction of pregnancy complications, including fetal growth restriction, gestational diabetes, hypertensive complications and premature births. Published clinical data about fetal growth restriction were only available to model Smarter Pregnancy, for which there was a 2.6% decreased occurrence of fetal growth restriction per couple. Indeed, the evidence for pre-conception nutrition associated with birth weight is compelling, with studies advocating “Mediterranean”-style diets high in fruit, vegetables, vegetable oil, fish, pasta, and rice as well as lower consumption of meat and potatoes [28,29]. For example, the size of the embryo represented by the crown-rump length is improved by an energy-rich nutritious diet (effect estimate 1.62, 95% CI 0.52-2.72; $P<.05$) [29]. Moreover, pre-conception diets with increased pre-conception omega-3 polyunsaturated fatty acid intake from fruit and vegetables was associated with improved

embryo morphology (linear regression coefficient β=0.6, *P*≤.05) [30]. Similarly, a "traditional Western" diet (high intake potatoes, mayonnaise and other fatty sauces, meat products, refined grains, sugar, and confectionary) was associated with an sperm DNA damage (linear regression coefficient β=13.25, *P*=.01), whereas a "health conscious" diet (high intakes of fruit, vegetables, fish and other seafood, whole grains and legumes) was inversely associated with sperm DNA damage (β=–2.81, *P*=.05) [31].

In contrast, published clinical data on gestational diabetes, hypertensive complications, and premature births were only available to model LIFEstyle, the pre-conception intervention to help obese women lose weight prior to and during early pregnancy. We modeled that implementation of the LIFEstyle program for obese women in the Netherlands seeking reproductive assistance may result in 4.4% lower chance of gestational diabetes, 3.8% lower chance of hypertensive complications, and 3.0% chance of premature birth. In subanalyses of the LIFEstyle study, women in the intervention group significantly reduced their consumption of sugary drinks and savory snacks, as well as increased their physical activity [32], and reduced the likelihood of developing metabolic syndrome [33]. A recent meta-analysis has shown that healthy diets (Mediterranean Diet, Dietary Approaches to Stop Hypertension diet and Alternate Healthy Eating Index diet) were associated with 15%-38% reduced relative risk of gestational diabetes, and compared with no physical activity, any prepregnancy or early pregnancy physical activity was associated with 30% and 21% reduced odds of gestational diabetes, respectively [34].

Our model also showed that lifestyle intervention programs are cost-effective to improve the chances of pregnancy [2]. It should be noted, however, that our estimated cost savings are conservative, as we did not include the cost of ovulation induction in our model. Moreover, we used conservative data regarding the number of subfertile men and women in the Netherlands. In our model, LIFEstyle has the greatest cost saving of €1163 (US $1380.18) per couple, which was similar to a cost analysis of the LIFEstyle study showed an overall saving of €1278 (US $1512.92) per couple [17].

Limitations

Inherent to the model-based approach using existing literature and expert opinion input data, we show results which may not completely reflect the real-world. Moreover, we did not model the costs for ovulation induction as it has a high success rate for anovulatory infertility. Although our model was based on couples living in the Netherlands, it nevertheless reflects couples seeking fertility assistance in other wealthy developed countries, such as those in Western Europe, North America, and Australia [35]. Published literature pertaining some pregnancy outcomes was not available, which necessitated assumptions being made. The costs of fertility treatment vary greatly depending on country and assisted reproductive technology procedures used, and the costs of some procedures are uncertain. We modeled the potential impact on reproductive and cost of fertility treatment rather than collecting real-world data. Although every effort was made to include evidence-based input parameters, the prevalence of subfertility among men and women living in the Netherlands may have been underestimated in our model. However, by using a conservative estimate of the number, the results of our model are also conservative, meaning that it is possible that there could be greater overall reductions in the number of IVF, ICSI, and IUI treatments required, fewer pregnancy complications, and larger cost savings than we have reported here. It is important to stress that the estimated cost savings are likely an underestimation, as we have not modeled more indirect cost savings that likely are achieved by reducing pregnancy complications, such as for instance, a more expedited return to the workforce for the mother, as well as improved health of the newborn which likely will lead to fewer health care expenditures in the years following birth. In addition, we have not included cost savings that can reasonably be expected as a result of improved lifestyle in the mother and father, likely leading to fewer health care expenditures as well as potentially a greater contribution to the workforce (and therefore taxation).

In our model, we included a comparison of the effects of smoking on IVF and ICSI outcomes in men. To date, there have been no published studies on pregnancy outcomes regarding the use of smoking cessation apps or programs in couples seeking to become pregnant, therefore, we have likely underestimated the effects of smoking cessation on pregnancy outcomes and the cost implications of smoking cessation in terms of reducing the need for fertility assistance and pregnancy complications. In addition, the effects of second-hand smoke have not been considered in our model.

Another limitation is that the costs of fertility treatment vary greatly depending on country and procedures required, and the costs of some procedures are uncertain. The type of fertility treatment required would depend on the cause of subfertility, which we have not included or addressed in our model. Further investigations are required to understand to what extent lifestyle modifications can reduce the risk of pregnancy complications, as well as affect direct outcomes of fertility treatment.

Comparisons With Prior Work

Our model input parameters were mostly based on good-quality evidence from published literature; however, as is inherently the case with hypothesis-based modeling, some assumptions were made. There have been 2 randomized controlled trials [9,10] that have investigated pre-conception lifestyle interventions to help obese women to lose weight. Although the study LIFEstyle [11] did not increase the healthy live birth rate, there was an increased rate of spontaneous pregnancies specifically among anovulatory women. A further exploratory analysis [13] suggested that a periconceptional decrease in BMI in obese subfertile women could lead to a decrease of the rates of hypertensive pregnancy complications and preterm birth; however, further randomized controlled trials are required to confirm these results. Similarly, in the second study, significantly more live births were achieved through spontaneous pregnancies in the weight reduction group (10.5%) than in the control group (2.6%; *P*=.009) [10]. Results from another randomized controlled trial, the UK Pregnancies Better Eating and Activity Trial (UPBEAT) [36], showed that specific dietary patterns in obese women in early pregnancy are linked to

gestational diabetes; however, the early pregnancy UPBEAT intervention did not reduce the incidence of gestational diabetes in its cohort [37]. Nevertheless, a secondary analysis of UPBEAT suggested that women randomized to the UPBEAT intervention had healthier metabolic profiles than those who received standard care [38]. As such, overweight or obese couples are likely to benefit from pre-conception lifestyle modifications of improved nutrition and physical activity prior to fertility treatment in an effort increase the chances of spontaneous conception and to help reduce pregnancy complications of gestational diabetes, hypertensive complications, and preterm birth. Improving lifestyle during the periconception period (defined as at least 14 weeks before conception) and the first 10 weeks of pregnancy is likely to help alleviate the risk of several adverse birth outcomes, such as congenital malformations, fetal growth restriction and babies born small or large for gestational age, as well as maternal pregnancy complications such as gestational diabetes, hypertensive disorders, and premature delivery [11-13,27].

Our data are also supported by a previous model-based cost analysis for Smarter Pregnancy used by 793 subfertile women undergoing IVF treatment [2]. This program resulted in 86 additional pregnancies and saved €270,000 (approximately US $319,630) compared to usual care after 2 IVF cycles, with an incremental cost-effectiveness ratio of €–3050 (95% CI €–3960 to €–540; or approximately US $–3611, 95% CI US $–4688 to US $–639) per additional pregnancy. The largest cost saving was from avoided IVF treatment costs. Sensitivity analyses showed that Smarter Pregnancy would need to increase the ongoing pregnancy rate by at least 51% after 2 IVF cycles for cost saving. Thus, Smarter Pregnancy is potentially cost saving for subfertile couples after their first IVF treatment.

There have been no published randomized controlled trials specifically on pre-conception interventions to quit smoking or to reduce or stop alcohol consumption. Nevertheless, Smarter Pregnancy has assessed pre-conception healthy nutrition and lifestyle, including tailored advice based on questionnaire responses regarding smoking and alcohol consumption in both couples seeking fertility treatment, and couples conceiving spontaneously [12]. Importantly, the Smarter Pregnancy intervention has been the only study to show a positive correlation between women whose partners also made positive lifestyle modifications and achieving pregnancy [20]. In addition, Smarter Pregnancy was not tailored for obese women and their partner, but these couples also appreciated the program very much and showed comparable effectiveness. There remains a lack of randomized controlled trials in this area as first highlighted by a Cochrane review nearly a decade ago [39].

Current evidence on the effectiveness of internet or app-based interventions is limited, and further investigation is needed in order to fully appraise their impact on fertility outcomes such as pregnancy rate, as well as pregnancy complications and newborn health. A recent systematic review [40] looked into feasibility, acceptability, and effectiveness of mobile health lifestyle and medical apps during pregnancy in high-income countries, which may be an appropriate way of offering lifestyle intervention support to subfertile couples in the future. Lifestyle apps that aimed to improve health behavior, reduce gestational weight gain, and for smoking cessation were generally effective [40]. As such, internet-based technologies such as Smarter Pregnancy have the potential to answer some of these questions as well as raise awareness around the importance of pre-conception care when trying to conceive [21]. Moreover, cognitive behavior therapy or psychological support during fertility treatment is suggested to lead to significantly more viable pregnancies than routine care [41]. As such, pre-conception lifestyle interventions, including medical apps, have the potential to increase ongoing pregnancy rates and birth outcomes in subfertile couples [40].

We do not yet know if various lifestyle modifications such as smoking, alcohol, and nutrition have different effects that depend on the socioeconomic status, educational level, or social background of the couple trying to conceive. Recent analysis of Smarter Pregnancy suggests that the program has been more effective in women living in nondeprived neighborhoods, who were, however, less likely to complete the 24 weeks of coaching than women who lived in deprived neighborhoods [42]. Although subfertile couples seeking fertility treatment are often intrinsically motivated to make positive lifestyle changes to improve the likelihood of a successful pregnancy, the socioeconomic status of couples may impact lifestyle as well as their ability to change lifestyle habits.

Conclusions

In summary, lifestyle is an important public health issue that has a significant and cumulative impact on fertility. Appropriate counseling could result in substantial reductions in the referrals for fertility investigations and treatments [27]. In order to maximize the pregnancy rates during fertility care, many subfertile couples could benefit from pre-conception lifestyle interventions delivered before fertility treatment. Although more research is needed regarding the use of internet- and app-based technologies for lifestyle and lifestyle interventions prior to fertility treatment, our results can be used to support subfertile couples, fertility care providers, and policy makers involved in public health to optimize clinical outcomes at affordable costs. Pre-conception lifestyle interventions are likely to be a cost-effective way of supporting subfertile couples trying to conceive. Further implementation of the app Smarter Pregnancy also designed for subfertile couples seeking fertility assistance is likely to be cost-effective and allow data on reproductive outcomes to be collected.

Acknowledgments

Christina Birch Olsen, an employee of Ferring International, conducted literature searches. Celia J Parkyn, PhD, provided medical writing and linguistic support during manuscript development, and was funded by Ferring BV. The fertility expert working group

advisory board meetings, the cost-effectiveness modeling of the lifestyle interventions and medical writing support was funded by Ferring BV.

Authors' Contributions

RS-T, JL, EC, and LS made substantial contributions to the conception and design of the study. EC, LS, and LV made substantial contributions to the literature search, the acquisition of data, and data analysis. RS-T, AH, HG, and JL provided additional data needed for the model. All authors also attended 3 advisory workshops to discuss fertility and lifestyle interventions. LS and LV developed the fertility outcomes and cost-effectiveness model. All authors made substantial contributions to the interpretation of data. All authors critically reviewed multiple versions of the manuscript during development, including a draft outline as well as full drafts. All authors approved the final version of the manuscript.

Conflicts of Interest

EC is employed by Ferring BV, which funded this research. LS is and LV was (during data collection, analysis, and manuscript preparation) employed by KPMG and are advisors to Ferring BV. AH, HG, AB, GD, MS, JS, and JL received payment from Ferring BV, for attending the fertility expert working group advisory board meetings. The department of Obstetrics and Gynecology at the University Medical Center Groningen Hanzeplein received an unrestricted educational grant from Ferring BV. RS-T contributed from the department of Obstetrics and Gynecology, Erasmus University Medical Center, and as consultant of eHealth Care Solutions, to one working group meeting for which she received one payment from Ferring BV. JL received unrestricted research grants from Ferring BV, ZonMw, and the Dutch Heart Association. He also received consultancy fees from the following companies: Danone, Euroscreen/Ogeda, Ferring BV, and Titus Healthcare.

Multimedia Appendix 1
Table S1: Input parameters for specific lifestyle interventions per year.
[DOCX File , 43 KB - mhealth_v8i11e19570_app1.docx]

Multimedia Appendix 2
Table S2: Input parameters – assisted reproductive technology costs, assumptions and fertility data from the Netherlands.
[DOCX File , 24 KB - mhealth_v8i11e19570_app2.docx]

References

1. Gormack AA, Peek JC, Derraik JGB, Gluckman PD, Young NL, Cutfield WS. Many women undergoing fertility treatment make poor lifestyle choices that may affect treatment outcome. Hum Reprod 2015 Jul;30(7):1617-1624. [doi: 10.1093/humrep/dev094] [Medline: 25924654]
2. Oostingh EC, Hall J, Koster MPH, Grace B, Jauniaux E, Steegers-Theunissen RPM. The impact of maternal lifestyle factors on periconception outcomes: a systematic review of observational studies. Reprod Biomed Online 2019 Jan;38(1):77-94. [doi: 10.1016/j.rbmo.2018.09.015] [Medline: 30424937]
3. Rossi BV, Abusief M, Missmer SA. Modifiable risk factors and infertility: what are the connections? Am J Lifestyle Med 2014;10(4):220-231 [FREE Full text] [doi: 10.1177/1559827614558020] [Medline: 27594813]
4. Heindel JJ, Vandenberg LN. Developmental origins of health and disease: a paradigm for understanding disease cause and prevention. Curr Opin Pediatr 2015 Apr;27(2):248-253 [FREE Full text] [doi: 10.1097/MOP.0000000000000191] [Medline: 25635586]
5. Hanson M, Gluckman P, Bustreo F. Obesity and the health of future generations. Lancet Diabetes Endocrinol 2016 Dec;4(12):966-967. [doi: 10.1016/S2213-8587(16)30098-5] [Medline: 27743979]
6. Sermondade N, Faure C, Fezeu L, Shayeb AG, Bonde JP, Jensen TK, et al. BMI in relation to sperm count: an updated systematic review and collaborative meta-analysis. Hum Reprod Update 2013;19(3):221-231 [FREE Full text] [doi: 10.1093/humupd/dms050] [Medline: 23242914]
7. Norman RJ, Mol BWJ. Successful weight loss interventions before in vitro fertilization: fat chance? Fertil Steril 2018 Sep;110(4):581-586. [doi: 10.1016/j.fertnstert.2018.05.029] [Medline: 30196941]
8. Best D, Avenell A, Bhattacharya S, Stadler G. New debate: is it time for infertility weight-loss programmes to be couple-based? Hum Reprod 2017 Dec 01;32(12):2359-2365. [doi: 10.1093/humrep/dex313] [Medline: 29040540]
9. Mutsaerts MAQ, van Oers AM, Groen H, Burggraaff JM, Kuchenbecker WKH, Perquin DAM, et al. Randomized trial of a lifestyle program in obese infertile women. N Engl J Med 2016 May 19;374(20):1942-1953. [doi: 10.1056/NEJMoa1505297] [Medline: 27192672]
10. Einarsson S, Bergh C, Friberg B, Pinborg A, Klajnbard A, Karlström P, et al. Weight reduction intervention for obese infertile women prior to IVF: a randomized controlled trial. Hum Reprod 2017 Aug 01;32(8):1621-1630. [doi: 10.1093/humrep/dex235] [Medline: 28854592]

11. van Oers AM, Groen H, Mutsaerts MAQ, Burggraaff JM, Kuchenbecker WKH, Perquin DAM, LIFEstyle study group. Effectiveness of lifestyle intervention in subgroups of obese infertile women: a subgroup analysis of a RCT. Hum Reprod 2016 Dec;31(12):2704-2713. [doi: 10.1093/humrep/dew252] [Medline: 27798042]
12. van Dijk MR, Koster MPH, Willemsen SP, Huijgen NA, Laven JSE, Steegers-Theunissen RPM. Healthy preconception nutrition and lifestyle using personalized mobile health coaching is associated with enhanced pregnancy chance. Reprod Biomed Online 2017 Oct;35(4):453-460. [doi: 10.1016/j.rbmo.2017.06.014] [Medline: 28688924]
13. van Oers AM, Mutsaerts MAQ, Burggraaff JM, Kuchenbecker WKH, Perquin DAM, Koks CAM, LIFEstyle study group. Association between periconceptional weight loss and maternal and neonatal outcomes in obese infertile women. PLoS One 2018;13(3):e0192670 [FREE Full text] [doi: 10.1371/journal.pone.0192670] [Medline: 29590118]
14. de Klerk C, Hunfeld JAM, Duivenvoorden HJ, den Outer MA, Fauser BCJM, Passchier J, et al. Effectiveness of a psychosocial counselling intervention for first-time IVF couples: a randomized controlled trial. Hum Reprod 2005 May;20(5):1333-1338. [doi: 10.1093/humrep/deh748] [Medline: 15650042]
15. Gameiro S, Boivin J, Dancet E, de Klerk C, Emery M, Lewis-Jones C, et al. ESHRE guideline: routine psychosocial care in infertility and medically assisted reproduction-a guide for fertility staff. Hum Reprod 2015 Nov;30(11):2476-2485. [doi: 10.1093/humrep/dev177] [Medline: 26345684]
16. Li J, Long L, Liu Y, He W, Li M. Effects of a mindfulness-based intervention on fertility quality of life and pregnancy rates among women subjected to first in vitro fertilization treatment. Behav Res Ther 2016 Feb;77:96-104. [doi: 10.1016/j.brat.2015.12.010] [Medline: 26742022]
17. van Oers AM, Mutsaerts MAQ, Burggraaff JM, Kuchenbecker WKH, Perquin DAM, Koks CAM, LIFEstyle study group. Cost-effectiveness analysis of lifestyle intervention in obese infertile women. Hum Reprod 2017 Dec 01;32(7):1418-1426. [doi: 10.1093/humrep/dex092] [Medline: 28486704]
18. Gordon JS, Armin J, Giacobbi P, Cunningham JK, Johnson T, Abbate K, et al. Development and evaluation of the See Me Smoke-Free multi-behavioral mHealth app for women smokers. Transl Behav Med 2017 Feb 02. [doi: 10.1007/s13142-017-0463-7] [Medline: 28155107]
19. Schmidt CA, Romine JK, Bell ML, Armin J, Gordon JS. User Participation and Engagement With the See Me Smoke-Free mHealth App: Prospective Feasibility Trial. JMIR Mhealth Uhealth 2017 Oct 09;5(10):e142 [FREE Full text] [doi: 10.2196/mhealth.7900] [Medline: 28993302]
20. van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth Platform for Pregnancy on Nutrition and Lifestyle of the Reproductive Population: A Survey. JMIR Mhealth Uhealth 2016 May 27;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197] [Medline: 27234926]
21. van Dijk MR, Koster MP, Rosman AN, Steegers-Theunissen RP. Opportunities of mHealth in Preconception Care: Preferences and Experiences of Patients and Health Care Providers and Other Involved Professionals. JMIR Mhealth Uhealth 2017 Aug 17;5(8):e123 [FREE Full text] [doi: 10.2196/mhealth.7834] [Medline: 28818820]
22. Hajizadeh Maleki B, Tartibian B, Chehrazi M. The effects of three different exercise modalities on markers of male reproduction in healthy subjects: a randomized controlled trial. Reproduction 2017 Feb;153(2):157-174. [doi: 10.1530/REP-16-0318] [Medline: 27920258]
23. Hajizadeh Maleki B, Tartibian B. Moderate aerobic exercise training for improving reproductive function in infertile patients: A randomized controlled trial. Cytokine 2017 Apr;92:55-67. [doi: 10.1016/j.cyto.2017.01.007] [Medline: 28092795]
24. Zitzmann M, Rolf C, Nordhoff V, Schräder G, Rickert-Föhring M, Gassner P, et al. Male smokers have a decreased success rate for in vitro fertilization and intracytoplasmic sperm injection. Fertil Steril 2003 Jun;79 Suppl 3:1550-1554. [doi: 10.1016/s0015-0282(03)00339-x] [Medline: 12801558]
25. Open data from the Dutch Health Authority. Nederlandse Zorgautoriteit. URL: http://www.opendisdata.nl [accessed 2018-08-13]
26. Dutch Agency for Statistics. StatLine (CBS). URL: https://opendata.cbs.nl/ [accessed 2019-11-09]
27. Twigt JM, Bolhuis MEC, Steegers EAP, Hammiche F, van IWG, Laven JSE, et al. The preconception diet is associated with the chance of ongoing pregnancy in women undergoing IVF/ICSI treatment. Hum Reprod 2012 Aug;27(8):2526-2531. [doi: 10.1093/humrep/des157] [Medline: 22593431]
28. Vujkovic M, de Vries JH, Lindemans J, Macklon NS, van der Spek PJ, Steegers EAP, et al. The preconception Mediterranean dietary pattern in couples undergoing in vitro fertilization/intracytoplasmic sperm injection treatment increases the chance of pregnancy. Fertil Steril 2010 Nov;94(6):2096-2101. [doi: 10.1016/j.fertnstert.2009.12.079] [Medline: 20189169]
29. Bouwland-Both MI, Steegers-Theunissen RPM, Vujkovic M, Lesaffre EMEH, Mook-Kanamori DO, Hofman A, et al. A periconceptional energy-rich dietary pattern is associated with early fetal growth: the Generation R study. BJOG 2013 Mar;120(4):435-445 [FREE Full text] [doi: 10.1111/1471-0528.12086] [Medline: 23194298]
30. Hammiche F, Vujkovic M, Wijburg W, de Vries JHM, Macklon NS, Laven JSE, et al. Increased preconception omega-3 polyunsaturated fatty acid intake improves embryo morphology. Fertil Steril 2011 Apr;95(5):1820-1823. [doi: 10.1016/j.fertnstert.2010.11.021] [Medline: 21130435]
31. Vujkovic M, de VJH, Dohle GR, Bonsel GJ, Lindemans J, Macklon NS, et al. Associations between dietary patterns and semen quality in men undergoing IVF/ICSI treatment. Hum Reprod 2009 Jun;24(6):1304-1312. [doi: 10.1093/humrep/dep024] [Medline: 19228759]

32. van Elten TM, Karsten MDA, Geelen A, van Oers AM, van Poppel MNM, Groen H, LIFEstyle study group. Effects of a preconception lifestyle intervention in obese infertile women on diet and physical activity; A secondary analysis of a randomized controlled trial. PLoS One 2018;13(11):e0206888 [FREE Full text] [doi: 10.1371/journal.pone.0206888] [Medline: 30403756]
33. van Dammen L, Wekker V, van Oers AM, Mutsaerts MAQ, Painter RC, Zwinderman AH, LIFEstyle study group. Effect of a lifestyle intervention in obese infertile women on cardiometabolic health and quality of life: A randomized controlled trial. PLoS One 2018;13(1):e0190662 [FREE Full text] [doi: 10.1371/journal.pone.0190662] [Medline: 29324776]
34. Mijatovic-Vukas J, Capling L, Cheng S, Stamatakis E, Louie J, Cheung NW, et al. Associations of Diet and Physical Activity with Risk for Gestational Diabetes Mellitus: A Systematic Review and Meta-Analysis. Nutrients 2018 May 30;10(6) [FREE Full text] [doi: 10.3390/nu10060698] [Medline: 29849003]
35. Noncommunicable Diseases Country Profiles, 2018. World Health Organization. URL: http://www.who.int/nmh/publications/ncd-profiles-2018/en/
36. Flynn AC, Seed PT, Patel N, Barr S, Bell R, Briley AL, UPBEAT consortium. Dietary patterns in obese pregnant women; influence of a behavioral intervention of diet and physical activity in the UPBEAT randomized controlled trial. Int J Behav Nutr Phys Act 2016 Nov 29;13(1):124 [FREE Full text] [doi: 10.1186/s12966-016-0450-2] [Medline: 27894316]
37. Poston L, Bell R, Croker H, Flynn AC, Godfrey KM, Goff L, UPBEAT Trial Consortium. Effect of a behavioural intervention in obese pregnant women (the UPBEAT study): a multicentre, randomised controlled trial. Lancet Diabetes Endocrinol 2015 Oct;3(10):767-777 [FREE Full text] [doi: 10.1016/S2213-8587(15)00227-2] [Medline: 26165396]
38. Mills HL, Patel N, White SL, Pasupathy D, Briley AL, Santos Ferreira DL, UPBEAT Consortium. The effect of a lifestyle intervention in obese pregnant women on gestational metabolic profiles: findings from the UK Pregnancies Better Eating and Activity Trial (UPBEAT) randomised controlled trial. BMC Med 2019 Jan 21;17(1):15 [FREE Full text] [doi: 10.1186/s12916-018-1248-7] [Medline: 30661507]
39. Anderson K, Norman RJ, Middleton P. Preconception lifestyle advice for people with subfertility. Cochrane Database Syst Rev 2010 Apr 14(4):CD008189. [doi: 10.1002/14651858.CD008189.pub2] [Medline: 20393968]
40. Overdijkink SB, Velu AV, Rosman AN, van BMD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
41. Domar AD, Clapp D, Slawsby EA, Dusek J, Kessel B, Freizinger M. Impact of group psychological interventions on pregnancy rates in infertile women. Fertil Steril 2000 Apr;73(4):805-811. [doi: 10.1016/s0015-0282(99)00493-8] [Medline: 10731544]
42. Gootjes DV, van Dijk MR, Koster MP, Willemsen SP, Steegers EA, Steegers-Theunissen RP. Neighborhood Deprivation and the Effectiveness of Mobile Health Coaching to Improve Periconceptional Nutrition and Lifestyle in Women: Survey in a Large Urban Municipality in the Netherlands. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11664 [FREE Full text] [doi: 10.2196/11664] [Medline: 30973345]

Abbreviations

BMI: body mass index
IUI: intrauterine insemination
IVF: in vitro fertilization
ICSI: intracytoplasmic sperm injection
UPBEAT: UK Pregnancies Better Eating And Activity Trial

Edited by G Eysenbach; submitted 23.04.20; peer-reviewed by A Rosman, R Poluru; comments to author 12.06.20; revised version received 25.07.20; accepted 16.08.20; published 23.11.20.

Please cite as:
Steegers-Theunissen R, Hoek A, Groen H, Bos A, van den Dool G, Schoonenberg M, Smeenk J, Creutzberg E, Vecht L, Starmans L, Laven J
Pre-Conception Interventions for Subfertile Couples Undergoing Assisted Reproductive Technology Treatment: Modeling Analysis
JMIR Mhealth Uhealth 2020;8(11):e19570
URL: https://mhealth.jmir.org/2020/11/e19570
doi:10.2196/19570
PMID:33226349

Original Paper

Effectiveness of an mHealth Intervention Combining a Smartphone App and Smart Band on Body Composition in an Overweight and Obese Population: Randomized Controlled Trial (EVIDENT 3 Study)

Cristina Lugones-Sanchez[1], NP; Maria Antonia Sanchez-Calavera[2], MD, PhD; Irene Repiso-Gento[3], MD; Esther G Adalia[4], MSc; J Ignacio Ramirez-Manent[5], MD, PhD; Cristina Agudo-Conde[1], MSc; Emiliano Rodriguez-Sanchez[6], MD, PhD; Manuel Angel Gomez-Marcos[6], MD, PhD; Jose I Recio-Rodriguez[7*], PhD; Luis Garcia-Ortiz[8*], MD, PhD; EVIDENT 3 Investigators[9]

[1]Institute of Biomedical Research of Salamanca (IBSAL), Primary Care Research Unit of Salamanca (APISAL), Health Service of Castilla y León (SACyL), Salamanca, Spain

[2]Institute for Health Research Aragón (IISA), Department of Internal Medicine, Psychiatry and Dermatology, University of Zaragoza, Zaragoza, Spain

[3]Valladolid Rural Health Center I. Health Service of Castilla y León (SACyL), Valladolid, Spain

[4]University of Castilla-La Mancha, Health and Social Research Center, Cuenca, Spain

[5]Calvià Primary Care Center, Health Service of Balear Islands, Balear Islands, Spain

[6]Institute of Biomedical Research of Salamanca (IBSAL), Primary Care Research Unit of Salamanca (APISAL), Health Service of Castilla y León (SACyL), Department of Medicine, University of Salamanca, Salamanca, Spain

[7]Institute of Biomedical Research of Salamanca (IBSAL), Primary Care Research Unit of Salamanca (APISAL), Health Service of Castilla y León (SACyL), Department of Nursing and Physiotherapy, University of Salamanca, Salamanca, Spain

[8]Institute of Biomedical Research of Salamanca (IBSAL), Primary Care Research Unit of Salamanca (APISAL), Health Service of Castilla y León (SACyL), Department of Biomedical and Diagnostic Sciences, University of Salamanca, Salamanca, Spain

[9]Spanish Research Network for Preventive Activities and Health Promotion in Primary Care (REDIAPP), Barcelona, Spain

[*]these authors contributed equally

Corresponding Author:
Cristina Lugones-Sanchez, NP
Institute of Biomedical Research of Salamanca (IBSAL)
Primary Care Research Unit of Salamanca (APISAL)
Health Service of Castilla y León (SACyL)
Av. Portugal 83, 2nd Fl.
Salamanca, 37005
Spain
Phone: 34 923 291100 ext 54750
Email: crislugsa@gmail.com

Abstract

Background: Mobile health (mHealth) is currently among the supporting elements that may contribute to an improvement in health markers by helping people adopt healthier lifestyles. mHealth interventions have been widely reported to achieve greater weight loss than other approaches, but their effect on body composition remains unclear.

Objective: This study aimed to assess the short-term (3 months) effectiveness of a mobile app and a smart band for losing weight and changing body composition in sedentary Spanish adults who are overweight or obese.

Methods: A randomized controlled, multicenter clinical trial was conducted involving the participation of 440 subjects from primary care centers, with 231 subjects in the intervention group (IG; counselling with smartphone app and smart band) and 209 in the control group (CG; counselling only). Both groups were counselled about healthy diet and physical activity. For the 3-month intervention period, the IG was trained to use a smartphone app that involved self-monitoring and tailored feedback, as well as a smart band that recorded daily physical activity (Mi Band 2, Xiaomi). Body composition was measured using the InBody 230 bioimpedance device (InBody Co., Ltd), and physical activity was measured using the International Physical Activity Questionnaire.

Results: The mHealth intervention produced a greater loss of body weight (–1.97 kg, 95% CI –2.39 to –1.54) relative to standard counselling at 3 months (–1.13 kg, 95% CI –1.56 to –0.69). Comparing groups, the IG achieved a weight loss of 0.84 kg more than the CG at 3 months. The IG showed a decrease in body fat mass (BFM; –1.84 kg, 95% CI –2.48 to –1.20), percentage of body fat (PBF; –1.22%, 95% CI –1.82% to 0.62%), and BMI (–0.77 kg/m^2, 95% CI –0.96 to 0.57). No significant changes were observed in any of these parameters in men; among women, there was a significant decrease in BMI in the IG compared with the CG. When subjects were grouped according to baseline BMI, the overweight group experienced a change in BFM of –1.18 kg (95% CI –2.30 to –0.06) and BMI of –0.47 kg/m^2 (95% CI –0.80 to –0.13), whereas the obese group only experienced a change in BMI of –0.53 kg/m^2 (95% CI –0.86 to –0.19). When the data were analyzed according to physical activity, the moderate-vigorous physical activity group showed significant changes in BFM of –1.03 kg (95% CI –1.74 to –0.33), PBF of –0.76% (95% CI –1.32% to –0.20%), and BMI of –0.5 kg/m^2 (95% CI –0.83 to –0.19).

Conclusions: The results from this multicenter, randomized controlled clinical trial study show that compared with standard counselling alone, adding a self-reported app and a smart band obtained beneficial results in terms of weight loss and a reduction in BFM and PBF in female subjects with a BMI less than 30 kg/m^2 and a moderate-vigorous physical activity level. Nevertheless, further studies are needed to ensure that this profile benefits more than others from this intervention and to investigate modifications of this intervention to achieve a global effect.

Trial Registration: Clinicaltrials.gov NCT03175614; https://clinicaltrials.gov/ct2/show/NCT03175614.

International Registered Report Identifier (IRRID): RR2-10.1097/MD.0000000000009633

(JMIR Mhealth Uhealth 2020;8(11):e21771) doi:10.2196/21771

KEYWORDS

diet records; mobile app; telemedicine; electric impedance; obesity; body fat distribution; weight control

Introduction

Obesity and overweight are considered a major public health issue. The prevalence of obesity has reached epidemic levels, with 650 million adults worldwide estimated to be obese in 2016 [1]. Additionally, more than one-half of the population in Europe is classified as overweight or obese [2]. In the case of Spain, the proportion of obese people may reach 58% by 2030 [3]. The association of obesity with cardiovascular diseases and type 2 diabetes [4], among other diseases, is well known, and obesity also exacerbates cardiovascular risk factors [5]. Thus, it contributes to an increase in the mortality rate worldwide, as a more frequent cause than underweight or malnutrition [6]. This situation makes the development and implementation of weight management interventions imperative.

Because of the complex nature of obesity, a multifactorial strategy is needed. The modification of lifestyles is the cornerstone of weight management, including diverse aspects, such as reduced energy intake, increased energy expenditure, exercise, and behavior change strategies [7,8]. Primary care providers (PCPs) play a critical role in recommending appropriate weight-loss strategies. Moreover, the positive effect of PCP advice on patient engagement in weight loss efforts has been demonstrated [9]. Unfortunately, there are some barriers to obesity management, such as the lack of tools or training [10]. Also, interventions are usually individual and face-to-face, which generates more demands by the patient, thereby increasing the burden on health care professionals and the health care system.

Mobile health (mHealth) could be an excellent strategy for PCPs to implement with their patients to help them maintain lifestyle changes. Information and communication technologies (ICTs) have the potential to standardize and improve the quality of treatment provided and increase resources for prevention activities [11]. They also allow PCPs to address barriers through enhancing self-monitoring of the patient by registering progress or symptoms, which could improve feedback communication and enable PCPs to spend less time gathering routine data and more time engaging with patients. This means of interaction might enhance treatment outcomes as well as improve follow-up of some chronic diseases [12] while optimizing PCP time and reducing costs [13]. Every year, thousands of mobile apps are developed with the purpose of improving lifestyles. To ensure that these tools are able to have a positive influence, more studies are needed because most apps available are suboptimal in quality, meaning that they have inadequate scientific coverage and accuracy of weight-related information [14]. Compared with usual practice, the use of ICTs in the primary care context might help patients to achieve significant weight loss [15], including patients who are socioeconomically disadvantaged [16], thereby increasing egalitarian access to treatment. However, further research is needed to determine the optimal use of technology in weight loss, since the inclusion of small sample sizes, and the variability in study designs, follow-up times, and interventions, may hinder replication and comparison of results [17], leading to unclear conclusions in this regard.

In relation to weight loss interventions on body composition, some studies have reported the effect of an energy-restricted high-protein diet combined with exercise on decreasing fat-free mass [18] and leisure-time exercise in reducing fat mass [19]. In recent years, some studies have provided important findings related to the feasibility of ICT interventions in this practical setting, such as the LEAN study [20] and the IDEA study [21]. In addition, a recent pilot study assessing a telenutrition weight loss intervention in primary care showed greater loss of weight and body fat in obese men compared with usual care [22]. These results spotlight the need for more research in this field in order

to achieve the optimal combination of health tools and the time needed to achieve changes.

However, previous studies have usually considered weight and BMI as the main outcomes. Although BMI is easy to obtain, it is an indirect measure of body composition and, therefore, less accurate than other measures [23] in estimating the distribution of body fat, resulting in misclassification of obesity. Recent studies have highlighted other useful measures involved in weight regulation, such as fat-free mass (FFM), body fat mass (BFM), and percentage of body fat (PBF), which could better explain body composition changes during weight interventions [24]. These variables are analyzed by bioelectrical impedance analysis (BIA), an indirect measure that uses multiple electrical currents through the body to estimate the percentage of different types of body tissue. Regarding the PBF, the correlation between BIA and the reference measure—dual-energy X-ray absorptiometry—was 0.88 for a healthy population [25], with a mean difference of –1.83 (SD 4.1%) for all subjects. In addition, BIA is the most cost-effective method of measuring body composition [26], making it a good alternative for its estimation.

Furthermore, it is important to determine whether these technologies can increase weight loss and modify body composition to clinically significant levels, which would show that ICTs could potentially be useful in tackling obesity. This study aims to assess the short-term effectiveness of a 3-month intervention that includes a smartphone app in combination with a smart band to lose weight and change body composition in sedentary Spanish adults who are overweight or obese.

Methods

Design and Scope

EVIDENT 3 is a randomized controlled, multicenter clinical trial with two parallel groups with a follow-up period of 12 months. The study was conducted in a primary care setting. The Primary Care Research Unit in Salamanca (APISAL) at the Biomedical Research Institute of Salamanca (IBSAL) coordinated the project in five primary care centers belonging to the Network for Preventive Activity and Health Promotion (REDIAPP) (Salamanca, Valladolid, Cuenca, Palma de Mallorca, and Zaragoza). Between June 2017 and November 2019, evaluations were made at baseline and at the 3-month follow-up visit. The results presented in this paper correspond to the short-term effect (3 months) of the intervention on body composition, considered one of the EVIDENT 3 study's secondary outcomes.

Study Population

The participants were selected by random sampling among the patients attending a consultation with their family doctor in each participating center. The inclusion criteria were age between 20 and 65 years, a BMI between 27.5 kg/m^2 and 40 kg/m^2, agreement to participate in the study, and signing the informed consent document. A detailed description of inclusion and exclusion criteria has been published in the study protocol [27]. To determine the effect of the intervention on body composition, an additional criterion was set: only subjects with both body composition measurements (at baseline and 3-month visit) assessed using the InBody 230 Body Composition Analyzer (InBody Co., Ltd) were included in the analysis. Hence, the study sample consisted of 440 subjects.

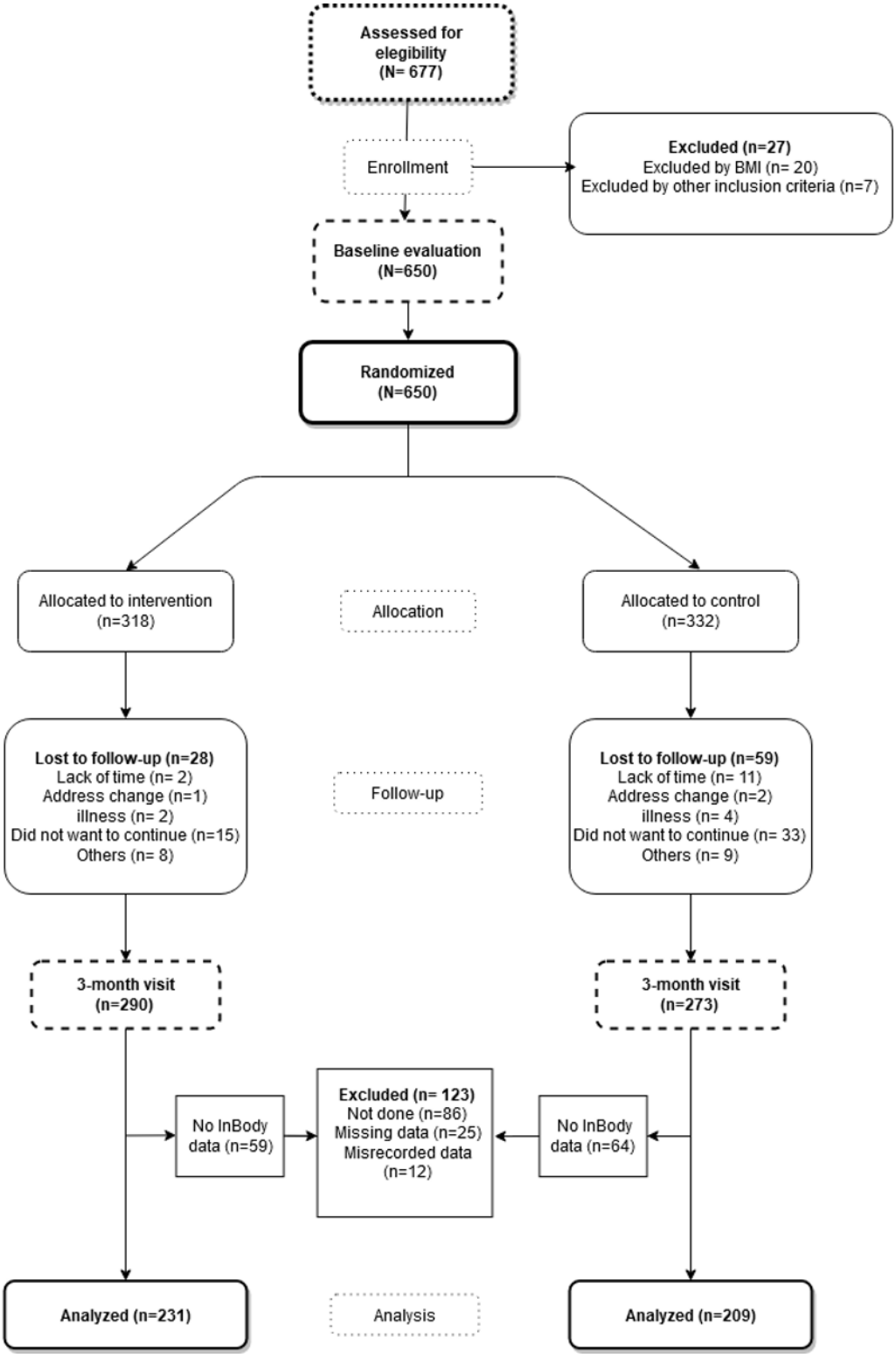
Sample Size

The sample size calculation was performed for the primary study endpoint of weight loss. Accepting an α risk of 0.05 and a β risk of 0.20, with an SD of 12 kg, and estimating from subjects from the EVIDENT 2 study [28], it was determined that 592 subjects would be needed (296 per group) to detect a decrease in weight of 3 kg or more [29] in the intervention group (IG) versus the control group (CG), taking into consideration a 15% loss of subjects at follow-up. This effect size represented a 3% to 5% difference between the groups, which was expected to produce clinically relevant health benefits [30]. There were 440 participants who completed the 3-month visit (IG, n=231; CG, n=209). Taking into account the sample size and a common SD of weight difference of 3.27 kg, the poststudy power to detect the 0.839 kg weight loss difference found between groups as significant was 77%.

Randomization

Participants were randomly assigned into two groups in a 1:1 ratio for the control group (CG) and intervention group (IG). Randomization was done after informed consent was obtained. The allocation sequence was generated through a standardized computer program (Epidat 4.2) by an independent researcher and concealed until the trial group was assigned (Figure 1). To minimize contamination between groups, the investigator who performed the intervention was different from the investigator who conducted the evaluation. The investigator who performed the data analysis was blinded to the subjects' groups. Due to the nature of the study, the subjects could not be blinded to the intervention.

Figure 1. Flow chart depicting study enrolment and completion.



Procedures

Each participant had to complete an initial visit and two follow-up visits, at 3 months and 12 months, after study inclusion. Baseline and follow-up data were collected by a research nurse. The IG completed an additional set of measurements at an appointment 7 days after baseline, with a different nurse performing the measurements, where the application was explained and the smart band given.

Outcome Measurements

The primary outcome was weight loss. Secondary outcomes included changes in some parameters of body composition. All outcomes were measured at baseline and 3 months after randomization.

Weight Loss

Body weight was measured to the nearest 0.1 kg, with the subject barefoot, wearing light clothing, and removing heavy

pocket items, using the portable InBody 230 Body Composition Analyzer (InBody Co., Ltd). Height was measured with the subject barefoot in a standing position using a portable scale and measurement system (Seca 222), and the average of 2 readings rounded to the nearest centimeter was recorded. BMI was calculated by dividing weight (in kg) by height squared (in m^2). Following the recommendations of the Spanish Society for the Study of Obesity (SEEDO) [31], waist circumference was measured in duplicate, using a flexible tape parallel to the floor, at the level of the midpoint between the last rib and the iliac crest, with the subject standing without clothing, after inspiration. Hip circumference was similarly measured at the level of the trochanters.

Body Composition

Body composition was determined by multifrequency BIA using an InBody 230 analyzer, with tetra-polar 8-point tactile electrodes that estimate total body water (TBW), dry lean mass, BFM, skeletal muscle mass (SMM), PBF, distribution of lean body mass, ratio of segmental lean mass, basal metabolic rate (BMR), and impedance of each body segment. This validated device [32] uses multiple currents at varying frequencies to provide precise body composition analysis without empirical estimation, increasing the reliability of the results.

The measurement was taken in the morning, before noon, with the subject barefoot, wearing light clothing, and standing upright for approximately 5 minutes before testing, with at least 2 hours of fasting and an empty bladder. These recommendations aim to measure body composition to the highest accuracy possible. The standing patient was required to wipe the palms of the hands, thumbs, and soles of the feet with the InBody tissue before placing them in the electrodes properly before testing. Individuals with medical implant devices such as pacemakers, essential support devices, or orthopedic prostheses, as well as pregnant women, could not be tested.

Clinically Relevant Measures

Data on the sociodemographic characteristics of the population including age, sex, education level, occupation, smoking history, and personal history of hypertension, dyslipidemia, and diabetes mellitus, as well as any active medical treatment, were collected. Smoking history was assessed through questions about the participant's smoking status (smoker or nonsmoker). We considered smokers to be subjects who currently smoked or who had stopped smoking less than 1 year before.

The short version of the International Physical Activity Questionnaire (IPAQ) [33] was used to measure activity subjectively. The IPAQ is a self-reported questionnaire that assesses physical activity performed at three intensity levels according to the energy expenditure estimated for each level: walking, moderate intensity, and vigorous intensity. For each level, participants reported frequencies such as days per week and average duration in minutes over the past week. This allowed the metabolic equivalents (METs) per minute per week to be calculated and subjects to be classified according to three activity levels: light, moderate, and vigorous.

Other variables were measured, including drug use, blood pressure, postprandial glucose, and biochemical parameters (total serum cholesterol, low-density lipoprotein-cholesterol, and high-density lipoprotein-cholesterol). A detailed description of how these variables were measured was published in the study protocol [27].

Intervention

Standard Counselling (CG and IG)

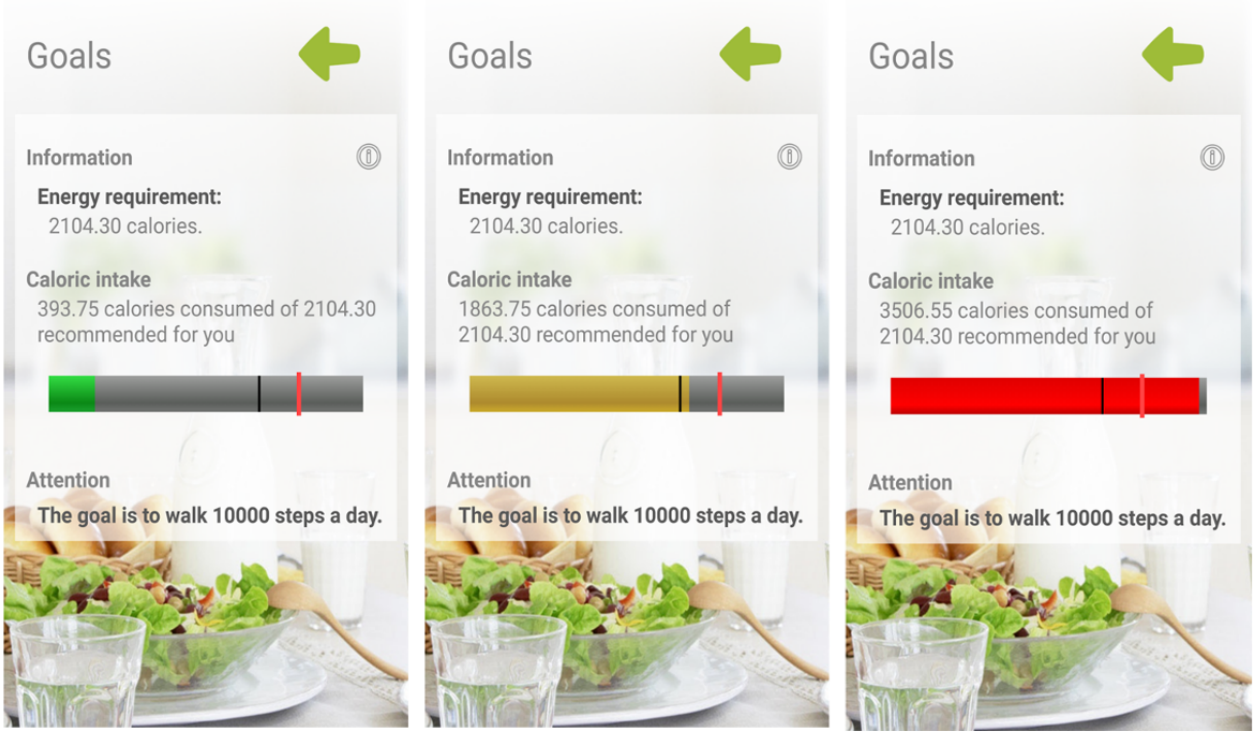
Both groups (CG and IG) received 5 minutes of counselling at the end of the baseline visit and prior to randomization. A trained nurse at each primary health center, who was not involved in other aspects of the study, gave advice on physical activity and healthy diet according to the current international recommendations for the general population. The health benefits of physical activity and the recommendation to complete at least 30 minutes of moderate activity 5 days a week, or 20 minutes of vigorous activity 3 days a week, were explained. Counselling on food was in compliance with the plate method [34].

Specific Intervention (IG)

The IG received the smartphone app and a smart band (Mi Band 2, Xiaomi) for 3 months, corresponding to the length of the intervention. Once the baseline visit was completed, another 15-minute visit was carried out 7 days later, at which the subjects were trained to use the device and the app (EVIDENT 3 Application [record entry no.00/2017/2438]), which was specifically designed for the study by CGB Computer Company and APISAL.

During the 15-minute visit, the app was configured with each participant's data (sex, age, weight, and height). It was designed to allow the dietary intake to be self-reported daily (Figure 2) and automatically record physical activity data from the smart band. Variables collected from the wearable device were number of steps taken, time of activity, kilometers traveled, and kilocalories expended. Subjects entered their food intake daily (divided into breakfast, midmorning snack, lunch, afternoon snack, and dinner) by selecting dishes and foods from the app menu and indicating the portion size. Thus, data collected comprised average energy intake (kcals), macronutrients and micronutrients (g/day), and time spent using the app (days). Food composition data were collected from the Spanish Food Composition Database [35] developed by RedBEDCA and AESAN. Once all of the daily information was collected, the app integrated the data to create personalized recommendations, based on the subjects' characteristics, and specific objectives and goals for weight loss. The subject was able to consult the app for these recommendations, as well as information about caloric intake changes and macronutrient distribution. At the 3-month visit, the devices were collected. All information generated by the app was duly analyzed and entered into the database.

Figure 2. EVIDENT app feedback information.



The behavioral strategies used in the intervention were those that enhance behavior changes toward healthier lifestyles. In this case, activities were meant to enhance self-efficacy, which is one of the most important determinants of behavior change [36,37], through self-monitoring, goal setting, and positive reinforcement. In order to avoid scheduling issues due to work shifts or other daily duties, participants were advised to use the app at the end of the day to register daily meals as well as to check physical activity information on the app. The smart band, worn at every moment, was set to congratulate the user when reaching 10,000 steps, following the general steps per day recommendation [38].

Blinding Strategy

The investigator carrying out the intervention with the IG was different from the person responsible for the assessment and standard counselling; both were kept blinded throughout the study, as was the investigator conducting data analysis. Due to the nature of the study, the subjects could not be blinded. To prevent contamination between groups, in the assessment visits (at 3 and 12 months), only the study variables were evaluated, but no advice or reinforcement could be given. In addition, the app was not available for download on the internet until the end of the study, so the CG could not make use of it in any way. During the follow-up visits, participants were instructed not to use other digital health technologies.

Ethical Considerations

The study was approved by the Clinical Research Ethics Committee of the Health Area of Salamanca on April 2016. All procedures were performed in accordance with the ethical standards of the institutional research committee and with the 2013 Declaration of Helsinki. All patients signed written informed consent documents prior to participation in the study. The trial was registered at ClinicalTrial.gov with identifier NCT03175614.

Statistical Analysis

General Analysis

The results were expressed as mean (SD) for quantitative variables and as frequency distributions for qualitative variables. The statistical normality was tested using the Kolmogorov-Smirnov test. Chi-square and Fisher tests were used to analyze the association between independent qualitative variables. Student *t* and Mann-Whitney U tests were used for the comparison of means between 2 independent groups. Pearson correlation and Spearman rho were used to evaluate the relationship between quantitative variables.

Analysis of Intervention Effect on Primary and Secondary Outcomes

To analyze the changes at 3 months after baseline in primary (weight loss) and secondary endpoints within the same group, we used the paired *t* test or McNemar test for quantitative or dichotomous variables, respectively. To analyze the effect of the intervention, we performed a multivariate analysis of variance of repeated measures, adjusted by the baseline value of each variable, in the follow-up for primary and secondary endpoints.

Analysis by Subgroups

We carried out subanalyses of the intervention effect on primary and secondary outcomes by sex (men and women), BMI at baseline (<30 kg/m^2 and ≥30 kg/m^2), and initial self-reported physical activity level (light and moderate-vigorous physical activity). Subanalyses were sufficiently powered (>65%) to detect differences in women, moderate-vigorous physical

activity, and overweight and obesity, but not in men or light physical activity.

The contrast in hypotheses established an α of .05. The data were analyzed using SPSS Statistics software (version 25.0; IBM Corp).

Results

Baseline Characteristics of the Participants and Follow-up

A total of 650 subjects fulfilled all of the inclusion criteria. They were included into the program and randomized to the IG or CG. Participant flow is presented in Figure 1. Testing at the 3-month visit was completed by 563 of 650 participants (86.6%).

In addition to the 87 subjects (13.4%) who dropped out during the study (Figure 1), 123 (123/650, 18.9%) subjects were excluded from the analysis. Exclusion requirements to perform the test were met in 86 subjects (86/650, 13.2%), so there were no measurements at any visit. In addition, the measurements of 37 subjects were not included due to incorrectly performed tests (25/650, 3.8%) or misrecorded data (12/650, 1.8%). Thus, 440 subjects (IG: n= 231; CG: n= 209) were included in the final analysis.

Both groups had a similar mean age—47.4 (SD 10.0) years in the IG and 48.8 (SD 9.2) years in the CG—and most participants were women (IG: 161/231, 69.7%; CG: 144/209, 68.9%) (Table 1). Most participants had middle or high school education or higher (206/231, 89.2% and 181/209, 86.6%) and a mean baseline BMI of 32.8 (SD 3.3) and 32.9 (SD 3.4) in the IG and CG, respectively. No differences in baseline characteristics were observed between the IG and CG.

Adherence to the smartphone app in the IG was calculated from app output data, showing that 129 of 231 (55.8%) subjects adhered sufficiently by using it for 60 days or more, 41 of 231 (17.8%) subjects used the app for 31 to 60 days, and 43 of 231 (18.6%) subjects entered data on 30 days or less. Two subjects (2/231, 0.9%) did not register any food intake information, and there were 16 corrupted files (16/231, 6.9%) from which no information was obtainable (Figure 3).

Regarding body composition variables, which were evaluated using BIA, no differences were found between the groups, with a mean weight of 89.7 kg (SD 13.1) in the IG and 90.7 kg (SD 13.9) in the CG. PBF, estimated using the InBody device, was 41.8% (SD 7.6%) and 42.1% (SD 6.4%) in the IG and CG, respectively. The main variables related to body composition are shown in Table 2.

In terms of self-reported physical activity, the IG had a total of 1263.6 METs/min/week and the CG had a total of 1353.3 METs/min/week, measured using the 7-day IPAQ, with no difference between them. At baseline, most of the sample showed a moderate physical activity level, in both the IG (50.2%) and the CG (51.2%).

Table 1. Baseline characteristics of the study population (N=440).

Baseline characteristics	Intervention group (n=231)	Control group (n=209)
Age (years), mean (SD)	47.4 (10.0)	48.8 (9.2)
Female sex, n (%)	161 (69.7)	144 (65.7)
Work situation, n (%)		
Works outside the home	170 (73.6)	157 (75.1)
Homemaker	20 (8.6)	14 (6.7)
Retired	14 (6.1)	13 (6.2)
Student	8 (3.5)	4 (1.9)
Unemployed	19 (8.2)	21 (10.1)
Educational level, n (%)		
University studies	97 (42.0)	88 (42.1)
Middle or high school	109 (47.2)	93 (44.5)
Elementary school	25 (10.8)	28 (13.4)
Smoking status, n (%)		
Nonsmoker	97 (42.0)	90 (43.1)
Smoker	46 (19.9)	44 (21.0)
Former smoker	88 (38.1)	75 (35.9)
Clinical variables, mean (SD)		
BMI (kg/m^2)	32.8 (3.3)	32.9 (3.4)
Waist circumference (cm)	105.9 (10.1)	107.1 (9.8)
Systolic blood pressure (mmHg)	118.3 (14.4)	119.8 (15.5)
Diastolic blood pressure (mmHg)	79.2 (8.7)	80.4 (9.9)
Total cholesterol (mg/dL)	199.7 (35.6)	201.7 (41.4)
Triglycerides (mg/dL)	127.3 (73.5)	127.3 (63.3)
Fasting plasma glucose (mg/dL)	92.5 (12.6)	94.3 (15.7)
Glycated hemoglobin (%)	5.4 (0.4)	5.5 (0.4)
Chronic diseases, n (%)		
Hypertension	57 (24.7)	72 (34.5)
Dyslipidemia	59 (25.5)	61 (29.2)
Diabetes	8 (3.5)	9 (4.3)
Medication use, n (%)		
Antihypertensive drugs	35 (15.2)	41 (19.6)
Lipid-lowering drugs	40 (17.3)	39 (18.7)
BMI classification, n (%)		
BMI≤30	60 (26.0)	50 (23.9)
BMI>30	171 (74.0)	159 (76.1)
Physical activity classification, n (%)		
Light physical activity	93 (40.3)	86 (41.2)
Moderate physical activity	116 (50.2)	107 (51.2)
Vigorous physical activity	22 (9.5)	16 (7.6)

Figure 3. Adherence to the EVIDENT smartphone app (number of days with a record in the app).

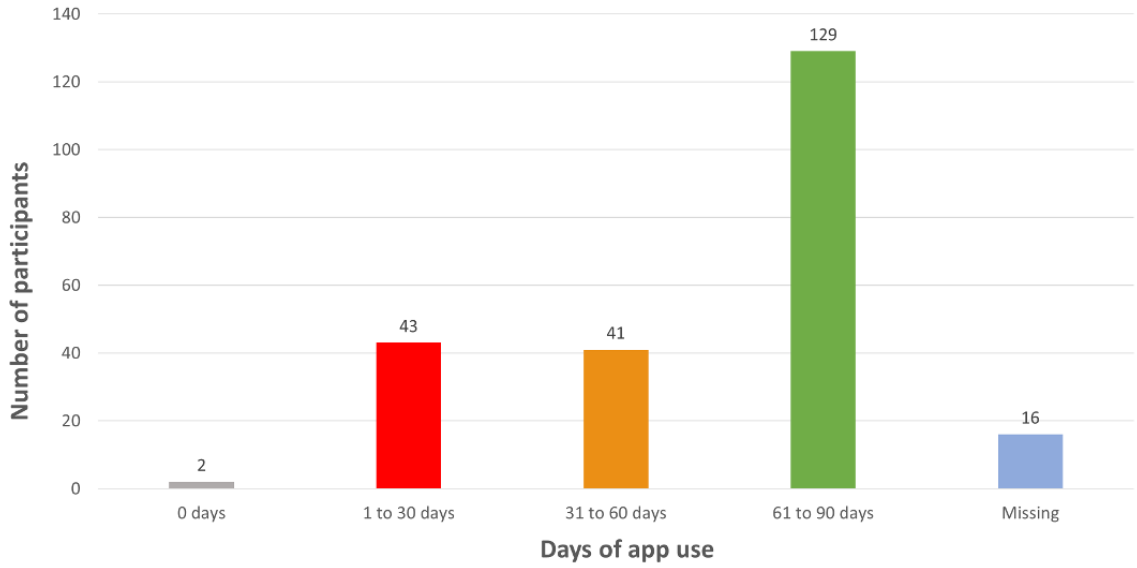


Table 2. Anthropometric and physical activity baseline data.

Variables	Intervention group (n=231), mean (SD)	Control group (n=209), mean (SD)	*P* value[a]
Body composition			
Weight (kg)	89.7 (13.1)	90.7 (13.9)	.43
Total body water (kg)	38.4 (7.9)	38.6 (7.8)	.73
Protein (kg)	10.3 (2.2)	10.4 (2.1)	.76
Minerals (kg)	3.6 (0.7)	3.7 (0.7)	.72
Body fat mass (kg)	37.4 (8.5)	38.1 (7.9)	.39
Fat-free mass (kg)	52.3 (10.8)	52.6 (10.6)	.74
Skeletal muscle mass (kg)	29.1 (6.6)	29.2 (6.4)	.77
BMI (kg/m^2)	32.7 (3.3)	32.9 (3.4)	.62
Body fat (%)	41.8 (7.6)	42.1 (6.4)	.68
Basal metabolic rate (kcal/day)	1499.4 (233.0)	1506.8 (229.3)	.74
Waist-to-hip ratio	1.0 (0.1)	1.0 (0.1)	.18
Physical activity			
METs[b] of intense activity	164.5 (258.9)	677.9 (1333.8)	.34
METs of moderate activity	218.9 (141.2)	656.6 (587.8)	.19
METs of light activity	880.1 (884.8)	901.6 (968.0)	.96
Total METs/minute/week	1263.6 (1285.0)	1353.3 (1723.2)	.88

[a]*P* value differences between intervention group and control group.

[b]METs: metabolic equivalents.

Changes in Body Weight During Study Period

The IG showed large changes in body weight (–1.97 kg, 95% CI –2.39 to –1.54) between baseline and 3 months, while the change was smaller (–1.13 kg, 95% CI –1.56 to –0.69) in the CG. Comparing groups, the IG achieved a weight loss of 0.84 kg more than the CG at 3 months (Table 3).

Analyzing by sex, there were no significant changes observed in body weight among men in the CG and IG. However, women in the IG had a significant weight loss of 1.37 kg (95% CI –2.03 to –0.71) compared with their CG counterparts.

Table 3. Effect of intervention on body composition variables for the total sample and by sex.

	Difference at 3 months		
Body composition variables	Intervention group (IG), mean (95% CI)	Control group (CG), mean (95% CI)	IG-CG, mean (95% CI)
All subjects			
Weight (kg)	–1.97 (–2.39 to –1.54)*	–1.13 (–1.56, –0.69)*	–0.84 (–1.45 to –0.23)*
TBW[a] (kg)	–0.04 (–0.42 to 0.34)	–0.01 (–0.39 to 0.37)	–0.03 (–0.57 to 0.50)
Protein (kg)	–0.04 (–0.17 to 0.08)	0.00 (–0.11 to 0.11)	–0.04 (–0.21 to 0.12)
Minerals (kg)	–0.04 (–0.11 to 0.03)	–0.01 (–0.04 to 0.02)	–0.03 (–0.11 to 0.05)
BFM[b] (kg)	–1.84 (–2.48 to –1.20)*	–1.11 (–1.69 to –0.53)*	–0.73 (–1.59 to 0.14)
FFM[c] (kg)	–0.13 (–0.63 to 0.38)	–0.01 (–0.53 to 0.50)	–0.11 (–0.83 to 0.60)
SMM[d] (kg)	–0.12 (–0.49 to 0.25)	0.02 (–0.31 to 0.35)	–0.14 (–0.64 to 0.36)
BMI (kg/m^2)	–0.77 (–0.96 to –0.57)*	–0.23 (–0.46 to 0.01)	–0.54 (–0.84 to –0.24)*
PBF[e] (%)	–1.22 (–1.82 to –0.62)*	–0.79 (–1.34 to –0.25)*	–0.42 (–1.24 to 0.39)
BMR[f] (kcal/day)	–2.63 (–13.49 to 8.23)	–0.30 (–11.37 to 10.77)	–2.34 (–17.82 to 13.15)
WHR[g]	–0.03 (–0.07 to 0.01)	–0.01 (–0.01 to 0.00)*	–0.02 (–0.07 to 0.02)
Men			
Weight (kg)	–1.70 (–2.54 to –0.85)*	–2.02 (–3.03 to –1.02)*	0.33 (–0.97 to 1.62)
TBW (kg)	0.06 (–0.37 to 0.50)	–0.17 (0.55 to 0.20)	0.24 (–0.33 to 0.81)
Protein (kg)	0.02 (–0.11 to 0.15)	–0.05 (–0.16 to 0.06)	0.07 (–0.10 to 0.24)
Minerals (kg)	0.01 (–0.01 to 0.04)	–0.02 (–0.06 to 0.03)	0.03 (–0.02 to 0.08)
BFM (kg)	–1.80 (–2.77 to –0.83)*	–1.79 (–2.51 to –1.07)*	–0.01 (–1.22 to 1.20)
FFM (kg)	0.10 (–0.47 to 0.68)	–0.23 (–0.74 to 0.27)	0.34 (–0.43 to 1.11)
SMM (kg)	0.11 (–0.27 to 0.49)	–0.11 (–0.43 to 0.20)	0.22 (–0.27 to 0.71)
BMI (kg/m^2)	–0.58 (–0.87 to –0.29)*	–0.55 (–0.92 to –0.18)*	–0.03 (–0.49 to 0.43)
PBF (%)	–1.48 (–2.39 to –0.57)*	–1.14 (–1.60 to –0.68)*	–0.34 (–1.37 to 0.69)
BMR (kcal/day)	2.35 (–10.11 to 14.81)	–5.05 (–15.98 to 5.89)	7.40 (–9.13 to 23.93)
WHR	–0.01 (–0.02 to 0.00)**	–0.01 (–0.03 to 0.00)*	0.00 (–0.01 to 0.02)
Women			
Weight (kg)	–2.08 (–2.58 to –1.59)*	–0.71 (–1.14 to –0.28)*	–1.37 (–2.03 to –0.71)*
TBW (kg)	–0.09 (–0.61 to 0.43)	0.07 (–0.46 to 0.60)	–0.16 (–0.89 to 0.58)
Protein (kg)	–0.07 (–0.24 to 0.10)	0.02 (–0.13 to 0.17)	–0.09 (–0.32 to 0.14)
Minerals (kg)	–0.06 (–0.17 to 0.04)	0.00 (–0.04 to 0.04)	–0.06 (–0.18 to 0.06)
BFM (kg)	–1.86 (–2.68 to –1.04)*	–0.80 (–1.58 to –0.02)**	–1.06 (–2.19 to 0.08)
FFM (kg)	–0.23 (–0.91 to 0.45)	0.09 (–0.63 to 0.80)	–0.32 (–1.30 to 0.67)
SMM (kg)	–0.22 (–0.73 to 0.29)	0.08 (–0.38 to 0.54)	–0.30 (–0.99 to 0.39)
BMI (kg/m^2)	–0.89 (–1.13 to –0.66)*	–0.17 (–0.39 to 0.05)	–0.72 (–1.05 to –0.40)*
PBF (%)	–1.10 (–1.87 to –0.33)*	–0.63 (–1.41 to 0.14)	–0.47 (–1.56 to 0.62)
BMR (kcal/day)	–4.84 (–19.59 to 9.90)	1.90 (–13.55 to 17.34)	–6.74 (–28.02 to 14.54)
WHR	–0.04 (–0.10 to 0.03)	–0.01 (–0.01 to 0.00)	–0.03 (–0.10 to 0.04)

[a]TBW: total body water.

[b]BFM: body fat mass.

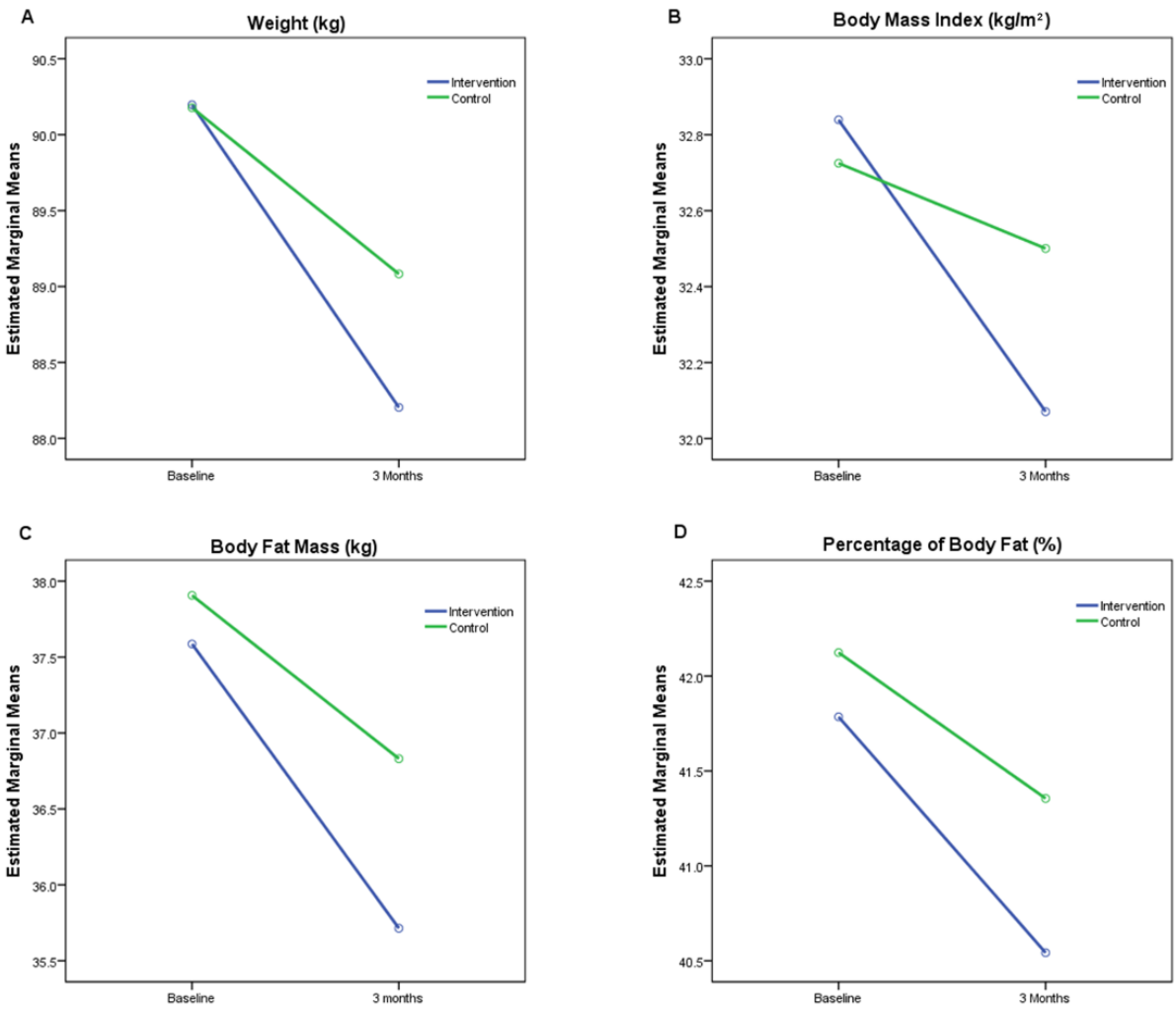
[c]FFM: fat-free mass.

[d]SMM: skeletal muscle mass.
[e]PBF: percentage of body fat.
[f]BMR: basal metabolic rate.
[g]WHR: waist-to-hip ratio.
*P<0.01.
**P<0.05.

Changes in Body Composition After Intervention

The IG showed a decrease in body composition variables (Figure 4), with a change of –1.84 kg (95% CI –2.48 to –1.20) in BFM, –1.22% (95% CI –0.96% to –0.57%) in PBF, and 0.77 kg/m^2 (95% CI –0.96 to –0.57) in BMI. A significant between-group difference was noted only in BMI (–0.54 kg/m^2, 95% CI –0.84 to –0.24).

Figure 4. Evolution over time of main body composition variables. (A) Weight. (B) BMI. (C) Body fat mass. (D) Percentage of body fat.



Analyzing by sex, men who received the intervention reduced BFM (–1.80 kg, 95% CI –2.77 to –0.83), PBF (–1.48%, 95% CI –2.39% to –0.57%), and BMI (–0.58 kg/m^2, 95% CI –0.87 to –0.29). Women and men in the IG achieved similar results, decreasing BFM (–1.86 kg, 95% CI –2.68 to –1.04), PBF (–0.89%, 95% CI –1.13% to –0.66%), and BMI (–0.89 kg/m^2, 95% CI –1.13 to –0.66). Although no significant changes were observed in any of these parameters in men, BMI was significantly reduced in women in the IG compared with women in the CG.

Changes in Weight and Body Composition by Baseline BMI Classification

Regarding BMI groups at baseline (Table 4), weight loss in the IG was greater in subjects with type I obesity (BMI≥30 kg/m^2) than in overweight (BMI<30 kg/m^2) subjects, with significant results in both cases, although compared with the CG, the differences were higher in the overweight group (–1.10 kg, 95% CI –2.01 to –0.18) than in the group with type I obesity (–0.77 kg, 95% CI –1.52 to –0.01).

Table 4. Effect of intervention on body composition according to BMI.

	Difference at 3 months		
	Intervention group (IG), mean (95% CI)	Control group (CG), mean (95% CI)	IG-CC, mean (95% CI)
BMI<30 kg/m^2			
Weight (kg)	–1.76 (–2.42 to –1.11)*	–0.67 (–1.31 to –0.03)**	–1.10 (–2.01 to –0.18)**
TBW[a] (kg)	0.31 (–0.16 to 0.77)	0.24 (–0.04 to 0.53)	0.06 (–0.51 to 0.63)
Protein (kg)	0.09 (–0.05 to 0.23)	0.06 (–0.02 to 0.14)	0.02 (–0.14 to 0.19)
Minerals (kg)	0.02 (–0.01 to 0.05)	0.02 (–0.01 to 0.04)	0.00 (–0.03 to 0.04)
BFM[b] (kg)	–2.17 (–3.10 to –1.25)*	–0.99 (–1.55 to –0.43)*	–1.18 (–2.30 to –0.06)**
FFM[c] (kg)	0.41 (–0.21 to 1.04)	0.33 (–0.06 to 0.71)	0.09 (–0.67 to 0.85)
SMM[d] (kg)	0.28 (–0.12 to 0.69)	0.21 (–0.02 to 0.44)	0.07 (–0.42 to 0.56)
BMI (kg/m^2)	–0.68 (–0.92 to –0.44)*	–0.22 (–0.46 to 0.02)	–0.47 (–0.80 to –0.13)*
PBF[e] (%)	–2.01 (–3.02 to –0.99)*	–0.99 (–1.49 to –0.49)*	–1.01 (–2.21 to 0.18)
BMR[f] (kcal/day)	8.95 (–4.46 to 22.36)	7.02 (–1.31 to 15.35)	1.93 (–14.45 to 18.31)
WHR[g]	–0.01 (–0.03 to 0.00)**	–0.01 (–0.01 to 0.00)**	–0.01 (–0.02 to 0.01)
BMI≥30 kg/m^2			
Weight (kg)	–2.04 (–2.57 to –1.50)*	–1.27 (–1.81 to –0.73)*	–0.77 (–1.52 to –0.01)**
TBW (kg)	–0.16 (–0.65 to 0.33)	–0.09 (–0.58 to 0.40)	–0.07 (–0.77 to 0.62)
Protein (kg)	–0.09 (–0.25 to 0.07)	–0.02 (–0.16 to 0.12)	–0.07 (–0.28 to 0.15)
Minerals (kg)	–0.06 (–0.16 to 0.04)	–0.01 (–0.05 to 0.02)	–0.05 (–0.15 to 0.06)
BFM (kg)	–1.72 (–2.52 to –0.92)*	–1.15 (–1.89 to –0.41)*	–0.57 (–1.67 to 0.52)
FFM (kg)	–0.32 (–0.96 to 0.33)	–0.12 (–0.78 to 0.54)	–0.20 (–1.12 to 0.73)
SMM (kg)	–0.26 (–0.74 to 0.22)	–0.04 (–0.47 to 0.38)	–0.22 (–0.86 to 0.43)
BMI (kg/m^2)	–0.84 (–1.07 to –0.60)*	–0.31 (–0.55 to –0.07)**	–0.53 (–0.86 to –0.19)**
PBF (%)	–0.94 (–1.67 to –0.22)**	–0.73 (–1.44 to –0.03)**	–0.21 (–1.22 to 0.80)
BMR (kcal/day)	–6.70 (–20.61 to 7.22)	–2.60 (–16.94 to 11.75)	–4.10 (–24.01 to 15.81)
WHR	–0.04 (–0.09 to 0.02)	–0.01 (–0.02 to 0.00)**	–0.03 (–0.09 to 0.04)

[a]TBW: total body water.

[b]BFM: body fat mass.

[c]FFM: fat-free mass.

[d]SMM: skeletal muscle mass.

[e]PBF: percentage of body fat.

[f]BMR: basal metabolic rate.

[g]WHR: waist-to-hip ratio.

*P<0.01.

**P<0.05.

In terms of body composition variables, both the CG and the IG showed reductions in BFM, PBF, and waist-to-hip ratio (WHR), with the reductions being greater in the IG. Comparing these groups, the biggest reductions were seen in the overweight group, with reductions in BFM (–1.18 kg, 95% CI –2.30 to –0.06) and BMI (–0.47 kg/m^2, 95% CI –0.80 to –0.13), whereas the group with type I obesity only decreased BMI (–0.53 kg/m^2, 95% CI –0.86 to –0.19). We observed no significant between-group differences in other study variables.

Changes in Weight and Body Composition by Baseline Self-Reported Physical Activity

When the data were analyzed according to the baseline physical activity level measured using the IPAQ (Table 5), there was a decrease in body composition variables in both the light physical activity and the moderate-vigorous physical activity groups within the IG. Participants in the light physical activity group lost similar weight (–1.99 kg, 95% CI –2.74 to –1.24) as those in the moderate-vigorous physical activity group (–1.95 kg, 95% CI –2.46 to –1.43). However, only the moderate-vigorous

physical activity group achieved a significant net loss compared with the CG (–1.03 kg, 95% CI –1.76 to –0.29).

In addition, the IG decreased body composition variables, showing reductions in the moderate-vigorous physical activity group such as –1.89 kg (95% CI –2.36 to –1.42) in BFM, –1.34% (95% CI –1.70% to –0.97%) in PBF, and –0.76 kg/m^2 (95% CI –0.95 to –0.57) in BMI, whereas the light physical activity group showed reductions of –1.77 kg (95% CI –3.21 to –0.33) in BFM and –0.85 kg/m^2 (95% CI –1.21 to –0.49) in BMI. Comparing these results with their counterparts in the CG, we found that only BMI (–0.51 kg/m^2, 95% CI –0.97 to –0.06) showed a significant difference in the light physical activity group, while the moderate-vigorous physical activity group reduced BFM (–1.03 kg, 95% CI –1.74 to –0.33), PBF (–0.76%, 95% CI –1.32% to –0.20%), and BMI (–0.51 kg/m^2, 95% CI –0.83 to –0.19) significantly. Differences in other body composition variables were not found.

Table 5. Effect of intervention on body composition according to physical activity at baseline.

	Difference at 3 months		
	Intervention group (IG), mean (95% CI)	Control group (CG), mean (95% CI)	IG-CC, mean (95% CI)
LPA[a]			
Weight (kg)	–1.99 (–2.74 to –1.24)*	–1.42 (–2.17 to –0.66)*	–0.57 (–1.63 to 0.49)
TBW[b] (kg)	–0.02 (–0.93 to 0.89)	0.06 (–0.76 to 0.89)	–0.08 (–1.31 to 1.15)
Proteins (kg)	–0.09 (–0.39 to 0.20)	0.01 (–0.23 to 0.25)	–0.10 (–0.49 to 0.28)
Minerals (kg)	–0.10 (–0.28 to 0.07)	–0.02 (–0.07 to 0.04)	–0.09 (–0.28 to 0.10)
BFM[c] (kg)	–1.77 (–3.21 to –0.33)**	–1.48 (–2.67 to –0.29)**	–0.29 (–2.16 to 1.59)
FFM[d] (kg)	–0.23 (–1.42 to 0.97)	0.06 (–1.05 to 1.17)	–0.29 (–1.92 to 1.34)
SMM[e] (kg)	–0.29 (–1.19 to 0.60)	0.07 (–0.66 to 0.79)	–0.36 (–1.52 to 0.80)
BMI (kg/m^2)	–0.85 (–1.21 to –0.49)*	–0.34 (–0.62 to –0.06)**	–0.51 (–0.97 to –0.06)**
PBF[f] (%)	–1.04 (–2.44 to 0.35)	–1.11 (–2.30 to 0.08)	0.07 (–1.77 to 1.90)
BMR[g] (kcal/day)	–4.78 (–30.67 to 21.10)	1.44 (–22.61 to 25.49)	–6.23 (–41.48 to 29.02)
WHR[h]	–0.06 (–0.17 to 0.05)	–0.02 (–0.03 to 0.00)**	–0.05 (–0.16 to 0.07)
MVPA[i]			
Weight (kg)	–1.95 (–2.46 to –1.43)*	–0.92 (–1.45 to –0.39)*	–1.03 (–1.76 to –0.29)**
TBW (kg)	–0.06 (–0.25 to 0.14)	–0.06 (–0.36 to 0.24)	0.00 (–0.35 to 0.35)
Protein (kg)	–0.01 (–0.06 to 0.05)	–0.01 (–0.09 to 0.08)	0.00 (–0.01 to 0.10)
Minerals (kg)	0.00 (–0.01 to 0.02)	0.00 (–0.03 to 0.03)	0.01 (–0.03 to 0.04)
BFM (kg)	–1.89 (–2.36 to –1.42)*	–0.85 (–1.39 to –0.32)**	–1.03 (–1.74 to –0.33)**
FFM (kg)	–0.06 (–0.33 to 0.21)	–0.07 (–0.48 to 0.34)	0.01 (–0.47 to 0.48)
SMM (kg)	0.00 (–0.17 to 0.16)	–0.02 (–0.27 to 0.23)	0.01 (–0.28 to 0.31)
BMI (kg/m^2)	–0.76 (–0.95 to –0.57)*	–0.25 (–0.52 to 0.01)	–0.51 (–0.83 to –0.19)*
PBF (%)	–1.34 (–1.70 to –0.97)*	–0.57 (–1.01 to –0.13)**	–0.76 (–1.32 to –0.20)**
BMR (kcal/day)	–1.18 (–6.93 to 4.57)	–1.51 (–10.37 to 7.34)	0.33 (–9.96 to 10.62)
WHR	–0.01 (–0.01 to 0.00)	0.00 (–0.01 to 0.00)	0.00 (–0.01 to 0.00)

[a]LPA: light physical activity.
[b]TBW: total body water.
[c]BFM: body fat mass.
[d]FFM: fat-free mass.
[e]SMM: skeletal muscle mass.
[f]PBF: percentage of body fat.
[g]BMR: basal metabolic rate.
[h]WHR: waist-to-hip ratio.
[i]MVPA: moderate-vigorous physical activity.
*P<.01.
**P<.05.

Discussion

Principal Findings

This study showed that the combined use of a mobile app and a smart band for 3 months, plus brief counselling at the start of the intervention, achieved a slight decrease in weight and BMI but not in other body composition variables. However, subanalyses by BMI, self-reported physical activity, and sex showed a greater decrease in variables such as BMI, WHR, BFM, and PBF in the IG with counselling than in the CG with counselling alone. Adding mHealth as a way of coaching and promoting healthy lifestyles in obese individuals may enhance weight loss outcomes at 3 months. More specifically, the intervention might be more effective in overweight women with

moderate physical activity, given that this group experienced greater reductions in weight and body composition variables.

This study offers relevant insight into the effect of mobile apps combined with wearable devices, such as an activity-tracking bracelet, on changing body composition with a large sample size. In recent years, interest in the effects of mHealth on body composition using BIA has increased, giving rise to research such as the TALENT study [39], in which an intensive, web-based lifestyle intervention (Individual Health Management) showed promising results, achieving a mean loss of approximately 10% of the baseline weight and a reduction in BMI, BFM, PBF, and waist circumference at 12 months. Even though the exercise intervention was on a web-based program, and no wearable devices were used, the results are in line with the results of this study. However, the IDEA study [21], in which one of the study groups was provided with wearable technology with a web-based interface for 24 months, did not find differences in body composition. Moreover, the study sample comprised only young adults, and it is not possible to generalize these results to other populations.

Taking studies with mobile apps into account, these results agree with other similar studies assessing the short-term effect of the Noom app (Noom Inc), a commercialized app that provides lifestyle-related logs, mainly food intake and exercise. The Noom app has been studied in combination with human coaching for 8 weeks [40] and alone for 15 and 52 weeks [41], achieving statistically significant decreases in weight, BMI, waist circumference, BFM, and PBF in both sexes. Nonetheless, none of these studies included a proper control group, requiring further research with high-quality methodology, which was the purpose of this study.

In terms of physical activity, a recent meta-analysis [42] demonstrated that a wearable technology intervention duration of more than 12 weeks was significantly more efficient than an intervention of fewer than 12 weeks in terms of BMI outcomes, and a systematic review [43] suggested that an activity tracker combined with a weight loss program may provide superior short-term (less than 6 months) results in middle-aged or older adults. Similarly, an intervention with Fitbit wrist activity tracker (Fitbit Inc) in medical students showed a positive trend for PBF in overweight women and lean body mass in overweight men [44], and another study using an app with push notifications to enhance diet and physical activity showed greater weight loss and body fat loss in obese women [45]. These results are in line with the results of this study, where there was a trend in BFM and PBF to decrease more in the IG, although no significant reduction was observed. In these studies, as in our research, a larger effect on fat mass was observed in women, which may be explained by the influence of psychological determinants, as women are more interested in participating in nutritional interventions [46] because of a desire to lose weight [47]. Also, we obtained a lower participation rate in men than in women, following the trend of the majority of studies of weight management [48], which could have led to body fat differences not being found in the male group in this study. Furthermore, body composition varies depending on sex, as women usually have a larger body fat mass proportion, whereas men are more likely to show greater lean mass, making a larger decrease of fat tissue in interventions with women plausible. Additionally, Slentz et al [49] reported that low amount/moderate-intensity and low-amount/vigorous-intensity endurance training (activity equivalent to 12 miles per week of walking or jogging) were equally effective in reducing the PBF, BFM, and waist circumference in sedentary adults. This result is in line with our findings and may explain part of the improvement in body composition variables in women through increased physical exercise.

The current results show the potential benefit of a short-term mHealth intervention with a mobile app, a smart band, and brief counselling as a useful tool for modifying body composition in overweight and obese healthy people in a primary health context. These findings are clinically relevant for various reasons. First, being an mHealth intervention with no professional face-to-face sessions or follow-up implies that there might be a cost reduction to implementing it in public health programs, and thus, this could be more cost-effective than other approaches. Through this study, we have identified the potential target profile for this intervention: overweight women (aged 18 to 65 years) with moderate physical activity at baseline. However, the physical activity classification was made using the IPAQ, which implies some degree of subjectivity and inaccuracy due to self-reporting. Future studies should explore the classification by accelerometer or another objective source of baseline physical activity in an attempt to obtain similar results. Nevertheless, these findings could be useful for adapting the intervention to population groups, whereby characteristics of each group could be taken into account in increasing the usefulness of the mHealth intervention.

Second, weight and/or BMI cannot solely be used as an accurate indicator of health, since body composition information is also relevant. The analysis of body composition could shed light on this field, allowing us to differentiate between the metabolically healthy but overweight and those with normal weight but with a pathological state. These states can be related to body components (eg, leg fat or SMM), endocrine interactions between individual fat deposits and muscle mass, and/or inflammation [50]. This stratification may be necessary to optimize prevention and treatment strategies and could be measured directly through bioimpedance variables. Furthermore, it is well known that excess body fat and its distribution are associated with metabolic syndrome and insulin resistance [51]. For these reasons, interventions that can modify body composition, focusing on decreasing BFM and PBF and leading to improved health markers, should be a priority in national health policies. In this sense, it is important to implement other measures related to fat distribution in addition to weight or BMI in daily clinical practice for a better approach. Even though our study showed moderate reductions in weight, BFM, and PBF in the IG, current findings support that a reduction in whole-body fat mass could predict changes in cardiometabolic health indices when increasing physical activity [52], even when body weight remains stable. In addition, physical activity seems to have a cardioprotective effect in subjects with higher PBF [53]. Thus, implementing this intervention in daily clinical practice could reduce the cardiovascular risk in overweight and obese people and its associated long-term issues.

Finally, it is important to point out the main limitations of this study. Missing InBody data were greater than expected (123 subjects). The majority of data lost were due to participants not meeting requirements to carry out the measurement, mainly because of medical implant or essential support devices. Despite this fact, losses were similar in both the IG and the CG. Due to the nature of the intervention, this could not be blinded to the participants; however, a recent meta-epidemiological study [54] suggested that blinding is less important than often believed. The duration of the intervention was only 3 months, so we could not measure the sustainability and long-term effect of the intervention. Also, despite the advice provided to subjects at baseline and follow-up visits to avoid the use of other apps related to nutrition or physical activity, we cannot guarantee that other apps were not used.

Conclusion

The results of this multicenter, randomized clinical trial study showed that, compared with standard counselling alone, using a self-reported app and a smart band obtained beneficial results in weight loss in women and a reduction in BFM and PBF in subjects with a BMI less than 30 kg/m^2 and moderate-vigorous physical activity level. Further studies are needed to ensure that this profile benefits more than others from this intervention and to investigate modifications of this intervention to achieve a global effect.

Acknowledgments

This study was funded by the Spanish Ministry of Science and Innovation, Instituto de Salud Carlos III, and co-funded by the European Union (ERDF/ESF, "Investing in your future") (RD12/0005/0001, RD16/0007/0003, RD16/0007/0005, RD16/0007/0008, RD16/0007/0009 and PI16/00101, PI16/00952, PI16/00765, PI16/00659, PI16/00421, PI16/00170, FI17/00040). Gerencia Regional de Salud de Castilla y Leon (GRS 1277/B/16) also collaborated in the funding of this study. They played no role in the study design, data analysis, reporting results, or decision to submit the manuscript for publication.

The EVIDENT 3 Investigators Group comprised the following:

Unidad de Investigacion de Atencion Primaria de Salamanca (APISAL):

Luis Garcia Ortiz, Jose I. Recio Rodriguez, Cristina Lugones-Sanchez, Manuel A. Gomez-Marcos, Cristina Agudo-Conde, Rosario Alonso-Dominguez, Natalia Sanchez-Aguadero, Angela de Cabo-Laso, Carmela Rodriguez-Martin, Carmen Castaño-Sanchez, Benigna Sanchez-Salgado, Emiliano Rodriguez-Sanchez, Susana Gonzalez-Sanchez, Jesus Gonzalez-Sanchez, Maria C Patino-Alonso, Jose A. Maderuelo-Fernandez, Rafael Hipola-Muñoz, Leticia Gomez-Sanchez, Olaya Tamayo-Morales, Inés Llamas-Ramos.

Centro de Salud Torreramona de Zaragoza (Health Service of Aragón):

Natividad González-Viejo, José Félix Magdalena-Belio, Luis Otegui-Ilarduya, Francisco J Rubio-Galan, Cristina I Sauras-Yera, Amor Melguizo-Bejar, Maria J Gil-Train, Marta Iribarne-Ferrer, Olga Magdalena-González, Miguel A Lafuente-Ripolles, M Mar Martínez.

Centro de Salud Cuenca I (Health Service of Castilla-La Mancha):

Fernando Salcedo-Aguilar, Fructuoso Muelas-Herraiz, Maria A Molina-Morate, Amparo Pérez-Parra, Fernando Madero, Angel Garcia-Imbroda, Jose M Izquierdo, María L Monterde.

Universidad de Castilla-La Mancha (University of Castilla-La Mancha):

Vicente Rodriguez-Vizcaino, Alba Soriano-Cano, Diana Patricia Pozuelo-Carrascosa, Esther Galvez-Adalia, Alicia del Saz-Lara, Ana Díez-Fernandez, Celia Alvarez-Bueno, Ivan Cavero-Redondo.

Centro de Salud Sta Ponça de Palma de Mallorca (Health Service of Balear Islands):

José I Ramírez-Manent, José L Ferrer-Perelló, José E Romero-Palmer, Manuel Sarmiento-Cruz, Guillermo Artigues, Jitka Mudrychova, María Albaladejo-Blanco, Margarita I Moyá-Seguí, Cristina Vidal-Ribas, Patricia Lorente-Montalvo, Isabel Torrens-Darder, María M Torrens-Darder, Lucía Pascual-Calleja.

Centro de Salud San Pablo de Valladolid (Health Service of Castilla y León):

Maria J Álvarez-Miguel, Maria D de Arriba-Gómez, Maria Á Rodríguez-Fernández, Isabel Arranz-Hernando, Silvia Ramos-de la Torre, Amparo Arqueaga-Luengo, Maria E Moreno-Moreno, Agustina Marcos-García, Nora Manrique-Vinagre, Nieves Palomo-Blazquez, José L Montalvillo-Montalvillo, Maria E Fernández-Rodríguez, Alejandro González-Moro, Marta Santiago-Pastor, Maria I Pérez-Concejo, Aurora Rubio-Fernández.

Centro de Salud Casa del Barco de Valladolid (Health Service of Castilla y León):

Amparo Gomez-Arranz, Carmen Fernandez-Alonso, Daniel Rodriguez-Dominguez, Irene Repiso-Gento, Aventina de la Cal-de la Fuente, Rosa Aragon-Garcia, Miguel A Diez-Garcia, Elisa Ibañes-Jalon, Ines Castrillo-Sanz, Ana M Corcho-Castaño, Esther Jimenez-Lopez, Daniel Correa-Gonzalez, Lucia Barruso-Villafaina, Isabel Peña-Garcia, Dolores Escudero-Terron, Pilar

Mena-Martin, Rosario Fraile-Gomez, Alberto Alonso-Gomez, Pilar Urueña, Francisca Martinez-Bermejo, Concepción Hernandez-San Jose, Manuela Nuñez-Gomez, Patricia Sanz-Capdepont, Ana I Pazos-Revuelta, Sofia Perez-Niño, María E Junquera-del Pozo.

CGB Computer Company, Salamanca, Spain, contributed to the technical development of EVIDENT 3 application.

Authors' Contributions

LGO and JIRR contributed to the conception and design of the study. LGO had full access to all of the data in the study and takes responsibility for the integrity of data and the accuracy of data analysis. CLS, LGO, and JIRR contributed to the drafting of the paper and CLS had the primary responsibility for final content. MAGM, LGO, and JIRR contributed to the analysis and interpretation of the data. CLS, MAGM, ERS, JIRR, and LGO contributed to the critical review of the paper for important intellectual content. CLS, MASC, IRG, JIRM, EGA, and CAC were responsible for collection and assembly of data. All the authors read and approved the final manuscript. LGO and JIRR contributed equally to this work.

Conflicts of Interest

None declared.

References

1. World Health Organization. Obesity and overweight. URL: http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2020-02-15]
2. Marques A, Peralta M, Naia A, Loureiro N, de MMG. Prevalence of adult overweight and obesity in 20 European countries, 2014. Eur J Public Health 2018 Apr 01;28(2):295-300. [doi: 10.1093/eurpub/ckx143] [Medline: 29036436]
3. Hernáez Á, Zomeño MD, Dégano IR, Pérez-Fernández S, Goday A, Vila J, et al. Exceso de peso en España: situación actual, proyecciones para 2030 y sobrecoste directo estimado para el Sistema Nacional de Salud. Revista Española de Cardiología 2019 Nov;72(11):916-924. [doi: 10.1016/j.recesp.2018.07.009]
4. Garg SK, Maurer H, Reed K, Selagamsetty R. Diabetes and cancer: two diseases with obesity as a common risk factor. Diabetes Obes Metab 2014 Feb;16(2):97-110 [FREE Full text] [doi: 10.1111/dom.12124] [Medline: 23668396]
5. Lavie CJ, McAuley PA, Church TS, Milani RV, Blair SN. Obesity and cardiovascular diseases: implications regarding fitness, fatness, and severity in the obesity paradox. J Am Coll Cardiol 2014 Apr 15;63(14):1345-1354 [FREE Full text] [doi: 10.1016/j.jacc.2014.01.022] [Medline: 24530666]
6. World Health Organization. Obesity: Preventing and managing the global epidemic: report of a WHO consultation. Geneva: WHO Report Series; 2000.
7. Khaylis A, Yiaslas T, Bergstrom J, Gore-Felton C. A review of efficacious technology-based weight-loss interventions: five key components. Telemed J E Health 2010 Nov;16(9):931-938 [FREE Full text] [doi: 10.1089/tmj.2010.0065] [Medline: 21091286]
8. Wadden TA, Butryn ML, Wilson C. Lifestyle modification for the management of obesity. Gastroenterology 2007 May;132(6):2226-2238. [doi: 10.1053/j.gastro.2007.03.051] [Medline: 17498514]
9. Rose SA, Poynter PS, Anderson JW, Noar SM, Conigliaro J. Physician weight loss advice and patient weight loss behavior change: a literature review and meta-analysis of survey data. Int J Obes (Lond) 2013 Jan;37(1):118-128. [doi: 10.1038/ijo.2012.24] [Medline: 22450855]
10. Kahan SI. Practical Strategies for Engaging Individuals With Obesity in Primary Care. Mayo Clin Proc 2018 Mar;93(3):351-359 [FREE Full text] [doi: 10.1016/j.mayocp.2018.01.006] [Medline: 29502565]
11. Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: a systematic review of reviews. Int J Med Inform 2010 Nov;79(11):736-771. [doi: 10.1016/j.ijmedinf.2010.08.006] [Medline: 20884286]
12. Zhang W, Cheng B, Zhu W, Huang X, Shen C. Effect of Telemedicine on Quality of Care in Patients with Coexisting Hypertension and Diabetes: A Systematic Review and Meta-Analysis. Telemed J E Health 2020 Sep 23. [doi: 10.1089/tmj.2020.0122] [Medline: 32976084]
13. Budman SH, Portnoy D, Villapiano AJ. How to get technological innovation used in behavioral health care: Build it and they still might not come. Psychotherapy: Theory, Research, Practice, Training 2003;40(1-2):45-54. [doi: 10.1037/0033-3204.40.1-2.45]
14. Chen J, Cade JE, Allman-Farinelli M. The Most Popular Smartphone Apps for Weight Loss: A Quality Assessment. JMIR Mhealth Uhealth 2015;3(4):e104 [FREE Full text] [doi: 10.2196/mhealth.4334] [Medline: 26678569]
15. Levine DM, Savarimuthu S, Squires A, Nicholson J, Jay M. Technology-assisted weight loss interventions in primary care: a systematic review. J Gen Intern Med 2015 Jan;30(1):107-117. [doi: 10.1007/s11606-014-2987-6] [Medline: 25134692]
16. Bennett GG, Steinberg D, Askew S, Levine E, Foley P, Batch BC, et al. Effectiveness of an App and Provider Counseling for Obesity Treatment in Primary Care. Am J Prev Med 2018 Dec;55(6):777-786 [FREE Full text] [doi: 10.1016/j.amepre.2018.07.005] [Medline: 30361140]

17. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A Systematic Review of Application and Effectiveness of mHealth Interventions for Obesity and Diabetes Treatment and Self-Management. Adv Nutr 2017 May;8(3):449-462. [doi: 10.3945/an.116.014100] [Medline: 28507010]
18. Wycherley TP, Noakes M, Clifton PM, Cleanthous X, Keogh JB, Brinkworth GD. A high-protein diet with resistance exercise training improves weight loss and body composition in overweight and obese patients with type 2 diabetes. Diabetes Care 2010 May;33(5):969-976 [FREE Full text] [doi: 10.2337/dc09-1974] [Medline: 20150293]
19. Quist JS, Rosenkilde M, Petersen MB, Gram AS, Sjödin A, Stallknecht B. Effects of active commuting and leisure-time exercise on fat loss in women and men with overweight and obesity: a randomized controlled trial. Int J Obes (Lond) 2018 Mar;42(3):469-478. [doi: 10.1038/ijo.2017.253] [Medline: 28993707]
20. Newton RL, Carter LA, Johnson W, Zhang D, Larrivee S, Kennedy BM, et al. A Church-Based Weight Loss Intervention in African American Adults using Text Messages (LEAN Study): Cluster Randomized Controlled Trial. J Med Internet Res 2018 Aug 24;20(8):e256 [FREE Full text] [doi: 10.2196/jmir.9816] [Medline: 30143478]
21. Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, et al. Effect of Wearable Technology Combined With a Lifestyle Intervention on Long-term Weight Loss: The IDEA Randomized Clinical Trial. JAMA 2016 Sep 20;316(11):1161-1171. [doi: 10.1001/jama.2016.12858] [Medline: 27654602]
22. Ventura Marra M, Lilly C, Nelson K, Woofter D, Malone J. A Pilot Randomized Controlled Trial of a Telenutrition Weight Loss Intervention in Middle-Aged and Older Men with Multiple Risk Factors for Cardiovascular Disease. Nutrients 2019 Jan 22;11(2):229. [doi: 10.3390/nu11020229]
23. Romero-Corral A, Somers VK, Sierra-Johnson J, Thomas RJ, Collazo-Clavell ML, Korinek J, et al. Accuracy of body mass index in diagnosing obesity in the adult general population. Int J Obes (Lond) 2008 Jun;32(6):959-966 [FREE Full text] [doi: 10.1038/ijo.2008.11] [Medline: 18283284]
24. Müller MJ, Braun W, Enderle J, Bosy-Westphal A. Beyond BMI: Conceptual Issues Related to Overweight and Obese Patients. Obes Facts 2016;9(3):193-205 [FREE Full text] [doi: 10.1159/000445380] [Medline: 27286962]
25. Sun G, French CR, Martin GR, Younghusband B, Green RC, Xie Y, et al. Comparison of multifrequency bioelectrical impedance analysis with dual-energy X-ray absorptiometry for assessment of percentage body fat in a large, healthy population. Am J Clin Nutr 2005 Jan;81(1):74-78. [doi: 10.1093/ajcn/81.1.74] [Medline: 15640463]
26. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu C, et al. Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study. Circulation 2007 Jul 03;116(1):39-48. [doi: 10.1161/CIRCULATIONAHA.106.675355] [Medline: 17576866]
27. Recio-Rodriguez JI, Gómez-Marcos MA, Agudo-Conde C, Ramirez I, Gonzalez-Viejo N, Gomez-Arranz A, EVIDENT 3 investigators. EVIDENT 3 Study: A randomized, controlled clinical trial to reduce inactivity and caloric intake in sedentary and overweight or obese people using a smartphone application: Study protocol. Medicine (Baltimore) 2018 Jan;97(2):e9633 [FREE Full text] [doi: 10.1097/MD.0000000000009633] [Medline: 29480874]
28. Recio-Rodríguez JI, Martín-Cantera C, González-Viejo N, Gómez-Arranz A, Arietaleanizbeascoa MS, Schmolling-Guinovart Y, et al. Effectiveness of a smartphone application for improving healthy lifestyles, a randomized clinical trial (EVIDENT II): study protocol. BMC Public Health 2014;14:254 [FREE Full text] [doi: 10.1186/1471-2458-14-254] [Medline: 24628961]
29. Voils CI, Olsen MK, Gierisch JM, McVay MA, Grubber JM, Gaillard L, et al. Maintenance of Weight Loss After Initiation of Nutrition Training: A Randomized Trial. Ann Intern Med 2017 Apr 04;166(7):463-471 [FREE Full text] [doi: 10.7326/M16-2160] [Medline: 28241185]
30. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. Circulation 2014 Jun 24;129(25 Suppl 2):S102-S138 [FREE Full text] [doi: 10.1161/01.cir.0000437739.71477.ee] [Medline: 24222017]
31. Salas-Salvadó J, Rubio MA, Barbany M, Moreno B, Grupo Colaborativo de la SEEDO. [SEEDO 2007 Consensus for the evaluation of overweight and obesity and the establishment of therapeutic intervention criteria]. Med Clin (Barc) 2007 Feb 10;128(5):184-96; quiz 1 p following 200. [doi: 10.1016/s0025-7753(07)72531-9] [Medline: 17298782]
32. Karelis AD, Chamberland G, Aubertin-Leheudre M, Duval C, Ecological mobility in AgingParkinson (EMAP) group. Validation of a portable bioelectrical impedance analyzer for the assessment of body composition. Appl Physiol Nutr Metab 2013 Jan;38(1):27-32. [doi: 10.1139/apnm-2012-0129] [Medline: 23368825]
33. Román Viñas B, Ribas Barba L, Ngo J, Serra Majem L. [Validity of the international physical activity questionnaire in the Catalan population (Spain)]. Gac Sanit 2013;27(3):254-257 [FREE Full text] [doi: 10.1016/j.gaceta.2012.05.013] [Medline: 23103093]
34. Raidl M, Spain K, Lanting R, Lockard M, Johnson S, Spencer M, et al. The healthy diabetes plate. Prev Chronic Dis 2007 Jan;4(1):A12 [FREE Full text] [Medline: 17173720]
35. RedBedca, AESAN. BEDCA. Base de datos Española de composición de alimentos. URL: https://bedca.net/ [accessed 2020-02-15]
36. Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJF, Martin BW, et al. Correlates of physical activity: why are some people physically active and others not? Lancet 2012 Jul 21;380(9838):258-271. [doi: 10.1016/S0140-6736(12)60735-1] [Medline: 22818938]

37. Linde JA, Rothman AJ, Baldwin AS, Jeffery RW. The impact of self-efficacy on behavior change and weight change among overweight participants in a weight loss trial. Health Psychol 2006 May;25(3):282-291. [doi: 10.1037/0278-6133.25.3.282] [Medline: 16719599]
38. Tudor-Locke C, Bassett DR. How many steps/day are enough? Preliminary pedometer indices for public health. Sports Med 2004;34(1):1-8. [Medline: 14715035]
39. Melchart D, Löw P, Wühr E, Kehl V, Weidenhammer W. Effects of a tailored lifestyle self-management intervention (TALENT) study on weight reduction: a randomized controlled trial. Diabetes Metab Syndr Obes 2017;10:235-245 [FREE Full text] [doi: 10.2147/DMSO.S135572] [Medline: 28684917]
40. Kim Y, Oh B, Shin H. Effect of mHealth With Offline Antiobesity Treatment in a Community-Based Weight Management Program: Cross-Sectional Study. JMIR Mhealth Uhealth 2020 Jan 21;8(1):e13273 [FREE Full text] [doi: 10.2196/13273] [Medline: 31961335]
41. Toro-Ramos T, Lee D, Kim Y, Michaelides A, Oh TJ, Kim KM, et al. Effectiveness of a Smartphone Application for the Management of Metabolic Syndrome Components Focusing on Weight Loss: A Preliminary Study. Metab Syndr Relat Disord 2017 Dec;15(9):465-473. [doi: 10.1089/met.2017.0062] [Medline: 29035677]
42. Yen H, Chiu H. The effectiveness of wearable technologies as physical activity interventions in weight control: A systematic review and meta-analysis of randomized controlled trials. Obes Rev 2019 Oct;20(10):1485-1493. [doi: 10.1111/obr.12909] [Medline: 31342646]
43. Cheatham SW, Stull KR, Fantigrassi M, Motel I. The efficacy of wearable activity tracking technology as part of a weight loss program: a systematic review. J Sports Med Phys Fitness 2018 Apr;58(4):534-548. [doi: 10.23736/S0022-4707.17.07437-0] [Medline: 28488834]
44. DiFrancisco-Donoghue J, Jung M, Stangle A, Werner WG, Zwibel H, Happel P, et al. Utilizing wearable technology to increase physical activity in future physicians: A randomized trial. Prev Med Rep 2018 Dec;12:122-127 [FREE Full text] [doi: 10.1016/j.pmedr.2018.09.004] [Medline: 30234000]
45. Hernández-Reyes A, Cámara-Martos F, Molina Recio G, Molina-Luque R, Romero-Saldaña M, Moreno Rojas R. Push Notifications From a Mobile App to Improve the Body Composition of Overweight or Obese Women: Randomized Controlled Trial. JMIR Mhealth Uhealth 2020 Feb 12;8(2):e13747. [doi: 10.2196/13747]
46. French SA, Jeffery RW, Wing RR. Sex differences among participants in a weight-control program. Addict Behav 1994;19(2):147-158. [doi: 10.1016/0306-4603(94)90039-6] [Medline: 8036962]
47. Livingstone KM, Celis-Morales C, Navas-Carretero S, San-Cristobal R, O'Donovan CB, Forster H, et al. Profile of European adults interested in internet-based personalised nutrition: the Food4Me study. Eur J Nutr 2016 Mar;55(2):759-769. [doi: 10.1007/s00394-015-0897-y] [Medline: 25893715]
48. Pagoto SL, Schneider KL, Oleski JL, Luciani JM, Bodenlos JS, Whited MC. Male inclusion in randomized controlled trials of lifestyle weight loss interventions. Obesity (Silver Spring) 2012 Jun;20(6):1234-1239 [FREE Full text] [doi: 10.1038/oby.2011.140] [Medline: 21633403]
49. Slentz CA, Duscha BD, Johnson JL, Ketchum K, Aiken LB, Samsa GP, et al. Effects of the amount of exercise on body weight, body composition, and measures of central obesity: STRRIDE--a randomized controlled study. Arch Intern Med 2004 Jan 12;164(1):31-39. [doi: 10.1001/archinte.164.1.31] [Medline: 14718319]
50. Müller MJ, Lagerpusch M, Enderle J, Schautz B, Heller M, Bosy-Westphal A. Beyond the body mass index: tracking body composition in the pathogenesis of obesity and the metabolic syndrome. Obes Rev 2012 Dec;13 Suppl 2:6-13. [doi: 10.1111/j.1467-789X.2012.01033.x] [Medline: 23107255]
51. Expert Panel on Detection, Evaluation,Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). JAMA 2001 May 16;285(19):2486-2497. [doi: 10.1001/jama.285.19.2486] [Medline: 11368702]
52. Amankwaah AF, Hudson JL, Kim JE, Campbell WW. Reductions in whole-body fat mass but not increases in lean mass predict changes in cardiometabolic health indices with exercise training among weight-stable adults. Nutr Res 2019 Mar;63:63-69. [doi: 10.1016/j.nutres.2018.11.004] [Medline: 30824398]
53. Calling S, Hedblad B, Engström G, Berglund G, Janzon L. Effects of body fatness and physical activity on cardiovascular risk: risk prediction using the bioelectrical impedance method. Scand J Public Health 2006;34(6):568-575. [doi: 10.1080/14034940600595621] [Medline: 17132589]
54. Moustgaard H, Clayton GL, Jones HE, Boutron I, Jørgensen L, Laursen DRT, et al. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. BMJ 2020 Jan 21;368:l6802 [FREE Full text] [doi: 10.1136/bmj.l6802] [Medline: 31964641]

Abbreviations

BFM: body fat mass
BIA: bioelectrical impedance analysis
BMR: basal metabolic rate

FFM: fat-free mass
ICTs: information and communication technologies
IPAQ: International Physical Activity Questionnaire
METs: metabolic equivalents
mHealth: mobile health
PBF: percentage of body fat
PCP: primary care provider
SEEDO: Spanish Society for the Study of Obesity
SMM: skeletal muscle mass
TBW: total body water
WHR: waist-to-hip ratio

Edited by G Eysenbach, L Buis; submitted 25.06.20; peer-reviewed by R Newton, J Chen; comments to author 14.09.20; revised version received 20.10.20; accepted 28.10.20; published 26.11.20.

Please cite as:
Lugones-Sanchez C, Sanchez-Calavera MA, Repiso-Gento I, Adalia EG, Ramirez-Manent JI, Agudo-Conde C, Rodriguez-Sanchez E, Gomez-Marcos MA, Recio-Rodriguez JI, Garcia-Ortiz L, EVIDENT 3 Investigators
Effectiveness of an mHealth Intervention Combining a Smartphone App and Smart Band on Body Composition in an Overweight and Obese Population: Randomized Controlled Trial (EVIDENT 3 Study)
JMIR Mhealth Uhealth 2020;8(11):e21771
URL: http://mhealth.jmir.org/2020/11/e21771/
doi:10.2196/21771
PMID:33242020

Original Paper

A Smartphone-Based Approach to Screening for Sudden Sensorineural Hearing Loss: Cross-Sectional Validity Study

Heng-Yu Haley Lin[1*], MPH, MD; Yuan-Chia Chu[2,3,4*], PhD; Ying-Hui Lai[5], PhD; Hsiu-Lien Cheng[5,6], MS; Feipei Lai[4,7,8], PhD; Yen-Fu Cheng[6,9,10,11], MD, PhD; Wen-Huei Liao[6,10], MD, PhD

[1]Department of Medical Education, Taipei Veterans General Hospital, Taipei, Taiwan
[2]Information Management Office, Taipei Veterans General Hospital, Taipei, Taiwan
[3]Big Data Center, Taipei Veterans General Hospital, Taipei, Taiwan
[4]Graduate Institute of Biomedical Electronics & Bioinformatics, National Taiwan University, Taipei, Taiwan
[5]Department of Biomedical Engineering, National Yang-Ming University, Taipei, Taiwan
[6]Department of Otolaryngology-Head and Neck Surgery, Taipei Veterans General Hospital, Taipei, Taiwan
[7]Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan
[8]Department of Computer Science & Information Engineering, National Taiwan University, Taipei, Taiwan
[9]Department of Medical Research, Taipei Veterans General Hospital, Taipei, Taiwan
[10]School of Medicine, National Yang-Ming University, Taipei, Taiwan
[11]Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan
[*]these authors contributed equally

Corresponding Author:
Wen-Huei Liao, MD, PhD
Department of Otolaryngology-Head and Neck Surgery
Taipei Veterans General Hospital
No. 201, Section 2, Shipai Road, Beitou District
Taipei
Taiwan
Phone: 886 2 2875 7808 ext 7264
Email: whliaovictor@gmail.com

Abstract

Background: Sudden sensorineural hearing loss (SSNHL) is an otologic emergency that warrants urgent management. Pure-tone audiometry remains the gold standard for definitively diagnosing SSNHL. However, in clinical settings such as primary care practices and urgent care facilities, conventional pure-tone audiometry is often unavailable.

Objective: This study aimed to determine the correlation between hearing outcomes measured by conventional pure-tone audiometry and those measured by the proposed smartphone-based Ear Scale app and determine the diagnostic validity of the hearing scale differences between the two ears as obtained by the Ear Scale app for SSNHL.

Methods: This cross-sectional study included a cohort of 88 participants with possible SSNHL who were referred to an otolaryngology clinic or emergency department at a tertiary medical center in Taipei, Taiwan, between January 2018 and June 2019. All participants underwent hearing assessments with conventional pure-tone audiometry and the proposed smartphone-based Ear Scale app consecutively. The gold standard for diagnosing SSNHL was defined as the pure-tone average (PTA) difference between the two ears being ≥30 dB HL. The hearing results measured by the Ear Scale app were presented as 20 stratified hearing scales. The hearing scale difference between the two ears was estimated to detect SSNHL.

Results: The study sample comprised 88 adults with a mean age of 46 years, and 50% (44/88) were females. PTA measured by conventional pure-tone audiometry was strongly correlated with the hearing scale assessed by the Ear Scale app, with a Pearson correlation coefficient of .88 (95% CI .82-.92). The sensitivity of the 5–hearing scale difference (25 dB HL difference) between the impaired ear and the contralateral ear in diagnosing SSNHL was 95.5% (95% CI 87.5%-99.1%), with a specificity of 66.7% (95% CI 43.0%-85.4%).

Conclusions: Our findings suggest that the proposed smartphone-based Ear Scale app can be useful in the evaluation of SSNHL in clinical settings where conventional pure-tone audiometry is not available.

(JMIR Mhealth Uhealth 2020;8(11):e23047) doi:10.2196/23047

KEYWORDS

sudden sensorineural hearing loss; hearing test; telemedicine; mobile apps; pure tone; audiometry

Introduction

Sudden sensorineural hearing loss (SSNHL) is an otologic emergency that warrants urgent clinical visits and timely management. SSNHL is commonly defined as a sensorineural hearing loss of 30 or more decibels (dB) over at least 3 consecutive audiometric frequencies occurring within a 72-hour period [1]; it affects approximately 5 to 27 per 100,000 people annually, and its incidence is gradually increasing over time [1-4]. Although SSNHL can occur at any age, the peak incidence occurs among adults aged 45 to 64 years, which is the general age range of working individuals [5]. The typical manifestations of SSNHL include immediate or rapidly progressive hearing loss and, sometimes, hearing loss upon awakening [1]. However, many patients with SSNHL often initially experience only nonspecific symptoms, such as aural fullness or a sensation of a blocked ear, and fail to recognize a loss of hearing, which results in delayed evaluations and treatment [1]. Compounded with the effects of aging and associated symptoms such as dizziness and tinnitus, SSNHL significantly impacts individuals' general health and quality of life and causes a considerable health care burden [1,6]. Previous studies have identified possible prognostic factors for hearing recovery following SSNHL, including age, severity of hearing loss, duration of hearing loss, and delay in treatment [5,7,8]. As it is a potentially modifiable variable, shortening the time between onset of hearing loss and adequate intervention is a crucial step in improving posttreatment hearing outcomes and minimizing other negative health consequences associated with hearing loss [9-12].

Currently, pure-tone audiometry remains the gold standard for evaluations of SSNHL since it not only reflects the severity of hearing loss but also provides a baseline hearing status for the assessment of recovery [5,8]. Conventional pure-tone audiometry usually requires a standard soundproof booth and calibrated audiometer, is performed by a qualified audiologist, and takes approximately 10 to 20 minutes per patient to perform. Considering the strict requirements regarding equipment and hearing care professionals, the accessibility of timely hearing evaluations using conventional pure-tone audiometry can be limited, especially in primary care settings [13,14]. To address these challenges and optimize the use of hearing health care, the traditional model of hearing screening and health service delivery should be supplemented with more efficient and attainable approaches. For hearing care, a hybrid hearing clinic with both internet-based and in-person services has been implemented in prior research and has revealed high patient satisfaction [15]. In terms of hearing screening, innovative telemedicine tools such as computer-assisted hearing tests [16-19] and mobile phone–based devices [20-23] have been introduced and investigated.

The Hearing Scale Test (HST) is a novel hearing screening tool derived from consecutive hearing screening procedures and used to estimate the current hearing status of each ear; it is based on the concepts of the Landolt C vision test chart [24,25]. With stratified hearing scales that represent various sound levels and four of the main frequencies in speech perception, 0.5 kHz, 1 kHz, 2 kHz, and 4 kHz, the HST not only precisely reflects an individual's hearing status but also has a computer-based design that enables outcome monitoring and patient surveillance [24]. The HST has demonstrated satisfactory feasibility and accuracy for hearing screening programs in pediatric populations in prior studies [24,25]. A recent study that integrated the HST into a smartphone-based app (Ear Scale) reported remarkable validity for hearing screening among school-aged children [26]. Given the paucity of studies applying innovative mobile phone–based hearing measures to screen for SSNHL, the aim of our study was to determine the correlations of hearing outcomes measured by the proposed smartphone-based hearing screening app with those measured by traditional pure-tone audiometry. We sought to determine the diagnostic validity of the smartphone-based hearing screening approach and, additionally, to explore its role and value in the evaluation of individuals with potential SSNHL.

Methods

Study Design and Population

This cross-sectional study was conducted at a tertiary medical center in Taipei, Taiwan, from January 2018 to June 2019. The sample size needed to reach a power of 0.80 was 82. We recruited 88 adults with possible SSNHL who visited either an otolaryngology outpatient clinic or emergency department. The study was approved by the institutional review board of the Taipei Veterans General Hospital (2016-12-004BC). Investigators explained the research objectives and process, and written informed consent was obtained from all patients enrolled. Instructions regarding the screening procedures and operations were provided by the trained examiners prior to each hearing screening test.

Hearing Measurements

Conventional Pure-Tone Audiometric Assessments

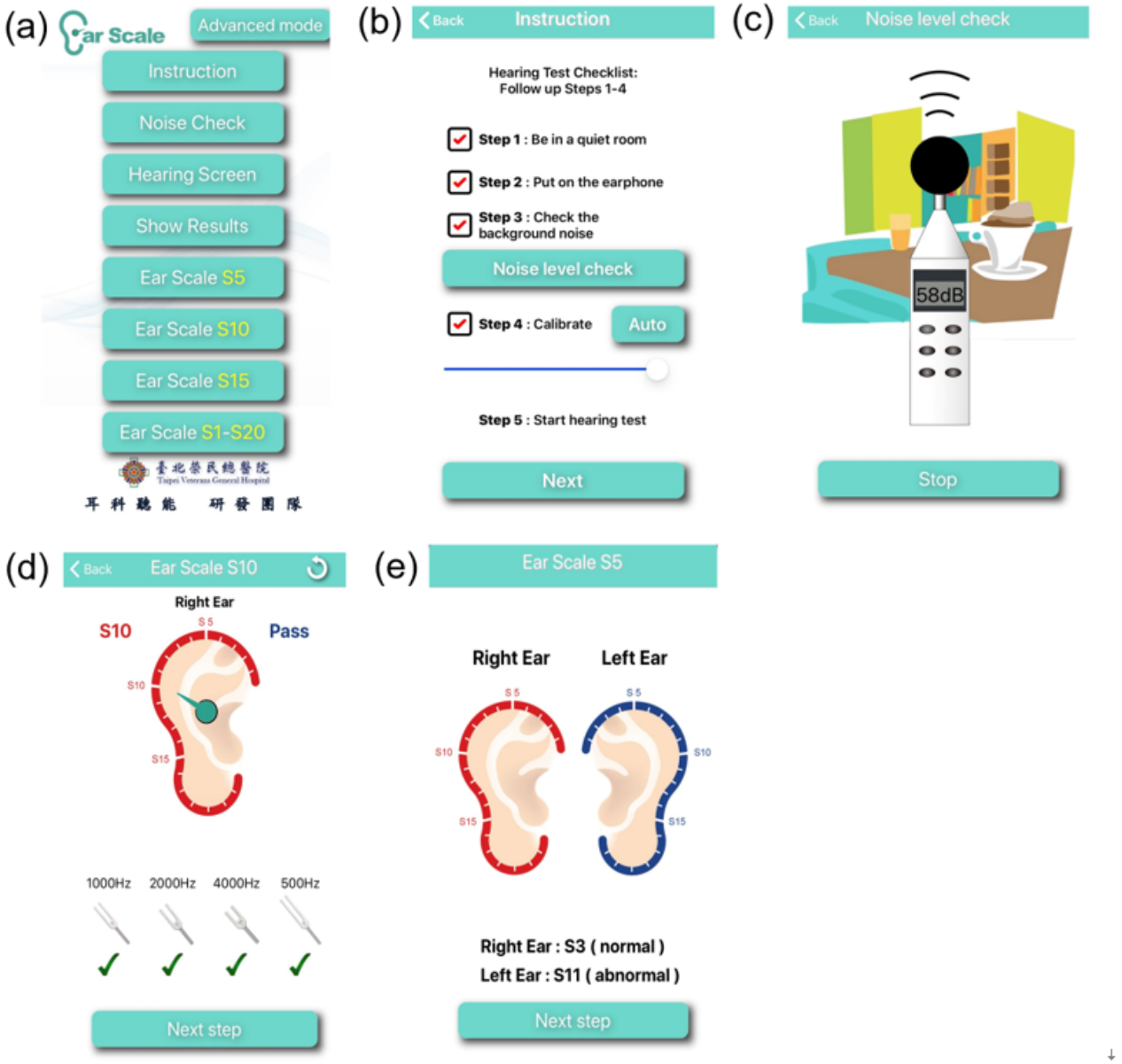
Pure-tone audiometry was administered by certified audiologists in the outpatient department. Otoscopy was performed to examine the clearness of the ear canal. Audiometric examinations were performed with a GSI 61 2-channel audiometer (Grason-Stadler Inc) in a soundproof booth. Standard clinical methods (modified Hughson-Westlake methods) were used to obtain pure-tone air conduction thresholds. To assess the reliability of the threshold measures, 1000 Hz was tested twice in each ear; participants with a >10 dB (dB) change between measures were considered unreliable. The pure-tone average (PTA) was calculated using air conduction thresholds at 0.5 kHz, 1 kHz, 2 kHz, and 4 kHz in each ear. Each individual's pretreatment hearing status measured by conventional pure-tone audiometry was categorized into the following 5 grades according to the modified Siegel criteria for

SSNHL [27]: grade 1 (PTA ≤25 dB HL), grade 2 (PTA 26-45 dB HL), grade 3 (PTA 46-75 dB HL), grade 4 (PTA 76-90 dB HL), and grade 5 (PTA >90 dB HL).

Smartphone-Based Hearing Screening App

The mobile devices used in this study were the iPhone 7 or iPhone 7 Plus (Apple Inc), with iOS software version 13.3.2. The iOS-based automated Ear Scale app (version 2.0) was integrated with the HST and used to measure the hearing statuses of both ears in the enrolled participants (Figure 1a). The items included in the hearing test checklist were assessed by examiners (Figure 1b). The patients were taught how to wear the headphones and click the response button when they heard the test tones. The headphones used throughout the examination were calibrated for Apple EarPods. The detailed calibration procedures are described in the next section. After the participants put on the headphones correctly, the background noise level was assessed immediately using the built-in function in the Ear Scale app to ensure that the ambient noise was less than 50 A-weighted decibels (Figure 1c). Last, the mobile device and headphones were calibrated and standardized before the HST was started (Figure 1b). The HST incorporated in the Ear Scale app was a novel hearing screening tool developed on the basis of consecutive hearing screening procedures to estimate the current hearing status of each ear [24,25]. The HST measured individuals' hearing status with respect to stratified hearing scales that represented sound intensity and 4 test frequencies (0.5 kHz, 1 kHz, 2 kHz, and 4 kHz). The adjacent scales differed from one another by 5 dB (Multimedia Appendix 1). The test tones lasted for 1.5 seconds, and the silent intervals lasted for 2 to 3 seconds [26,28]. The Ear Scale app started with hearing scale 5 (S5), which corresponded to 25 dB HL. The four test tones were automatically presented to patients in a fixed order of 1 kHz, 2 kHz, 4 kHz, and 0.5 kHz. The stimulus level of the pure tones descended to the next adjacent hearing scale only if the patient responded correctly to all tones (Figure 1d). The minimum audible hearing scale indicated the lowest pure-tone stimulus level at which the participant responded correctly to all four test tones, was shown at the end of each examination, and was saved to the devices (Figure 1e). The hearing scale difference between the impaired ear and the contralateral ear was determined and used for identifying patients with SSNHL (Figure 1e).

Figure 1. Screenshots of the Ear Scale app instructions for the subjects and the hearing test procedures.



iOS Automated Audiometry Calibration

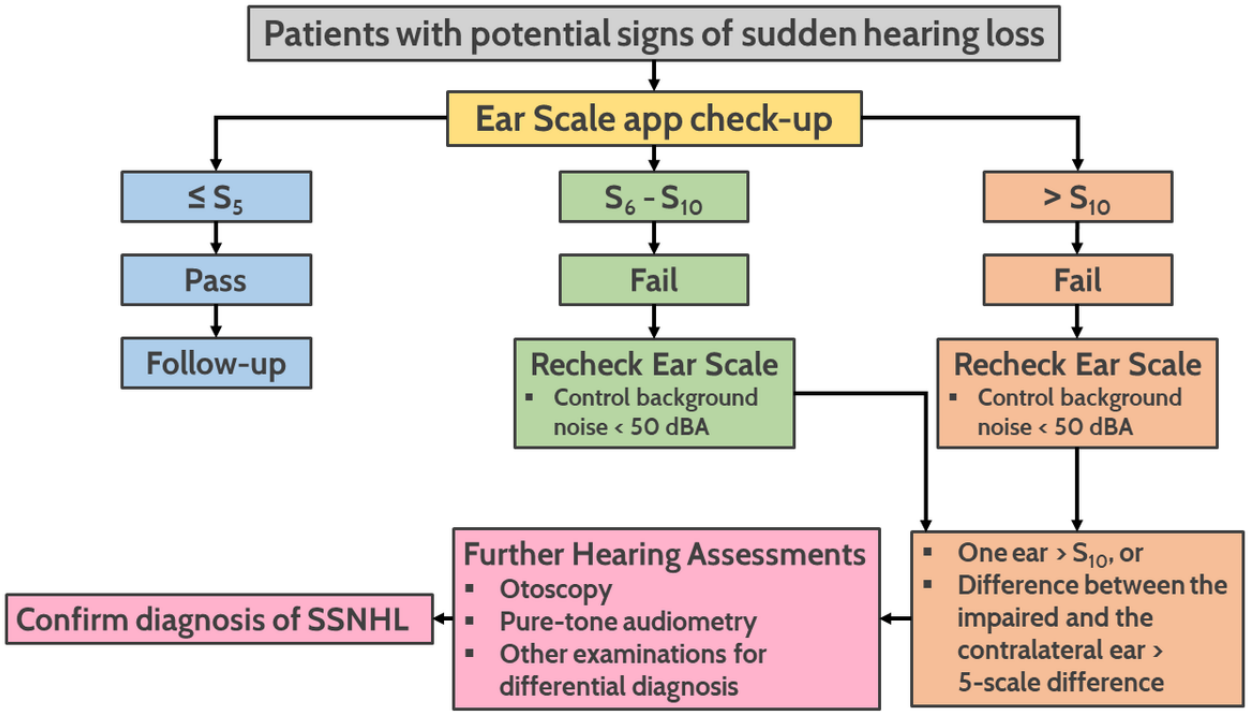
To calibrate the sound output of the iOS mobile devices to a hearing threshold level of zero at various frequencies, we applied the reference equivalent threshold sound pressure levels for Apple EarPods, which were reported in a prior study with consistent output across different EarPod pairs and between the right and left earphones and therefore could be applied to various Apple mobile devices with EarPods [29]. The Knowles Electronics Manikin for Acoustic Research (KEMAR) was initially invented in collaboration with the audiological industry for the use of hearing aid development. KEMAR meets the international standards that are specified by ISO, IEC as well as ANSI. To record the eardrum pressure and evaluate the sound quality, the EarPods were placed in the left and right pinna of a KEMAR manikin, which included a head and torso that had been designed specifically for anthropomorphic testing in the audiologic industry [30]. The microphones of the simulators and the electrical and acoustical measurement systems were calibrated using a 42AA Pistonphone (GRAS Sound & Vibration). The hearing thresholds were determined in an ascending order, as described in ISO 8253-1 [29], with a step size of 1 dB. The initial stimulus level was set to be 10 dB lower than the lowest subject response threshold, which was predetermined by conventional audiometry. Pure-tone stimuli at 0.25 kHz, 0.5 kHz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz were generated on the iOS mobile devices and delivered by the Apple EarPods. All the devices were standardized by setting the user-controllable volume to 100% of the maximum limit. A 2-down, 1-up adaptive staircase procedure was used to determine the final hearing threshold of each subject after 3 reversals [31]. The maximum output difference between the right and left EarPods was less than 1 dB, and the maximum output difference between the devices (iPhone 7 and iPhone 7 Plus) was less than 1.5 dB. The output levels of the EarPods were calibrated in units of dB sound pressure level when the volume of the Apple mobile device was set to the maximum. The output level (dBSPL) of

the pure tone at each test frequency was similar to that previously reported [28,29].

Hearing Screening Procedures

Figure 2 illustrates how the proposed Ear Scale app was used for hearing screening among patients enrolled in this study who had signs of possible sudden hearing loss. Participants underwent the Ear Scale app examination at presentation and were classified into 3 groups (≤S5, S6-S10, >S10) based on their test results. We then arranged comprehensive hearing assessments, including otoscopy, conventional pure-tone audiometry, and other examinations for those who had a hearing scale greater than S10 in one ear or asymmetrical hearing with hearing scale differences greater than 5-scale between the affected ear and the contralateral ear (Figure 2). Participants with bilateral sudden sensorineural hearing loss or conductive hearing loss were excluded from the study population.

Figure 2. Hearing screening procedures used in this study.



Statistical Analysis

Pearson correlation coefficients were estimated to investigate the correlation between PTA measured by conventional pure-tone audiometry and the hearing scale derived from the Ear Scale app using the HST, as well as the differences in the hearing results between the impaired ear and the contralateral ear of each individual. The corresponding PTA of each hearing scale group was demonstrated using a box plot. Analysis of variance was used to determine the difference in the mean PTA between each scale. Indicators of validity and the predictive value were estimated to determine the diagnostic accuracy of the HST for SSNHL compared with that of the gold standard pure-tone audiometric evaluation. Patients with a PTA difference between the two ears of at least 30 dB within a 72-hour period (ie, the diagnostic gold standard for identifying SSNHL that was used in this study), as assessed by conventional pure-tone audiometry, were considered positive for SSNHL. We then estimated the sensitivity, specificity, positive predictive value, and negative predictive value for diagnosing SSNHL on the basis of 3 hearing scale differences (5-scale difference, 6-scale difference, and 7-scale difference) between the two ears as measured by the Ear Scale app. Sensitivity was defined as the percentage of individuals with true SSNHL (ie, patients with PTA thresholds that met the diagnostic gold standard for the presence of SSNHL included in the American Academy of Otolaryngology-Head and Neck Surgery guidelines [1]) who were correctly identified as having SSNHL by the Ear Scale app. Positive predictive value was defined as the probability of true SSNHL being present among participants who were considered positive for SSNHL by the Ear Scale app. The significance tests for all analyses were 2-sided and included a type I error of .05. The power was set to be 0.80. The statistical software used was Stata 15 (StataCorp LLC).

Results

Baseline Characteristics of the Study Sample

This study included 88 adults with possible SSNHL who visited the emergency department or an otolaryngology clinic from January 2018 to June 2019; patients with bilateral or conductive hearing loss were excluded. The mean age of the study cohort was 46 years, and 50% (44/88) were females (Table 1). The average PTA of the cohort included in the analytic cohort was 67.1 dB HL (Table 1). The average hearing scale measured by the Ear Scale app was S17 (ie, 85 dB HL). Regarding the differences in the hearing results between the two ears, the mean PTA difference was 47.6 dB, whereas the average hearing scale difference (obtained from the Ear Scale app) was 9 hearing scales (ie, 45 dB difference; Table 1).

Table 1. Baseline characteristics of the study sample (n=88).

Variables	Values
Age in years, mean (SD)	46 (14.7)
Gender (female), n (%)	44 (50)
Pretreatment hearing grade of worst-hearing ear, n (%)	
Grade 1 (PTA[a] ≤25 dB HL[b])	7 (8)
Grade 2 (PTA 26-45 dB HL)	8 (9)
Grade 3 (PTA 46-75 dB HL)	43 (49)
Grade 4 (PTA 76-90 dB HL)	16 (18)
Grade 5 (PTA >90 dB HL)	14 (16)
PTA of worst-hearing ear, dB, mean (SD)	67.1 (24.9)
Average scale of worst-hearing ear, mean (SD)	17 (4.2)
Average PTA difference[c], dB, mean (SD)	47.6 (25.0)
Average scale difference[d], mean (SD)	9 (4.4)

[a]PTA: pure-tone average.

[b]dB HL: decibel hearing level.

[c]PTA difference = PTA of impaired ear – PTA of contralateral ear. [d]Hearing scale difference = hearing scale of impaired ear – hearing scale of contralateral ear.

Correlation Between the Pure-Tone Average and Hearing Scale

The Pearson correlation analyses revealed strong positive correlations between the PTA assessed by pure-tone audiometry and the hearing scale measured by the Ear Scale app as well as between the PTA differences and hearing scale differences between the two ears, with correlation coefficients of .88 (95% CI .82-.92) and .84 (95% CI .77-.90), respectively (Figure 3).

The association of the PTA and hearing scale differences between the two ears is presented in Figure 4. The mean PTA difference differed significantly across the hearing scale groups (P<.05).

Figure 3. Scatter plots demonstrating the (a) correlation between the pure-tone average (PTA) obtained by pure-tone audiometry (y-axis) and the hearing scale measured by the Ear Scale app (x-axis) and (b) correlation between the PTA differences and hearing scale differences between the impaired and contralateral ears.

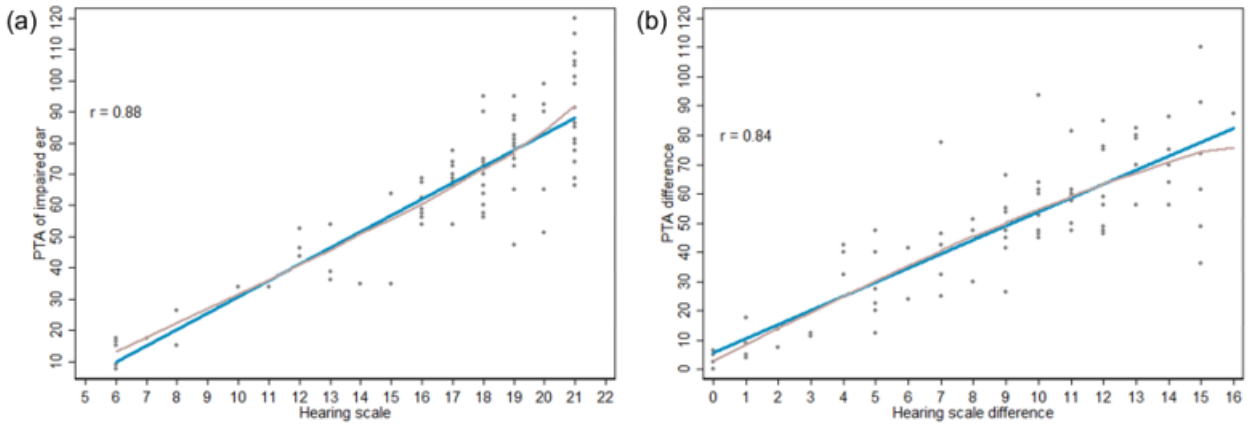
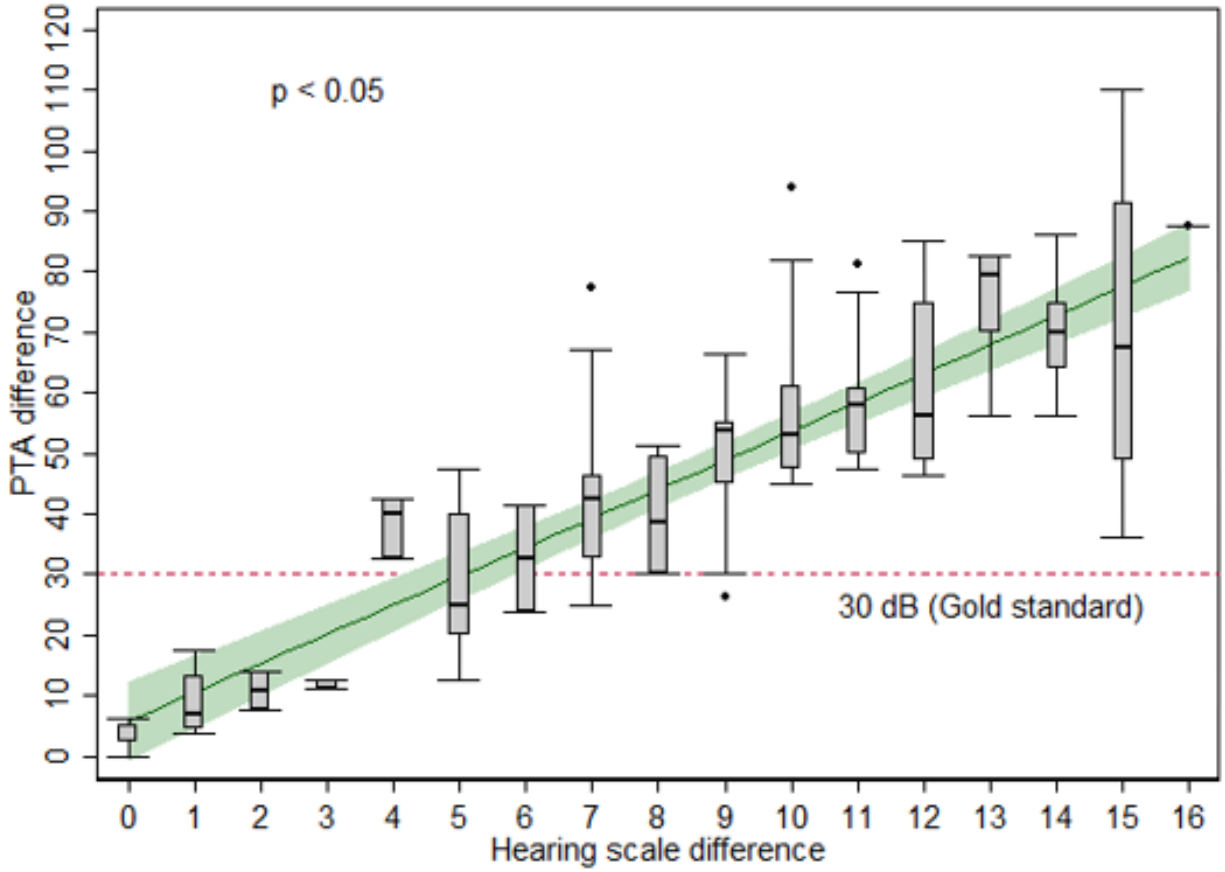


Figure 4. Box plot of the pure-tone average (PTA) difference (y-axis) in relation to the hearing scale difference (x-axis) between the impaired and contralateral ears. The green line depicts the best-fitted mean PTA difference in relation to the hearing scale difference for the linear regression, and the green area represents the 95% confidence interval of the model (P<.05, significant differences were found between each hearing scale difference group). The dashed line represents PTA differences of 30 dB (ie, diagnostic gold standard for detecting sudden sensorineural hearing loss in this study).



Validity of the Ear Scale App in Diagnosing Sudden Sensorineural Hearing Loss

The diagnostic gold standard for SSNHL used in our study was a PTA difference between the impaired ear and the contralateral ear of ≥30 dB. The indicators of validity of the Ear Scale app and the cutoff values for the hearing scale differences are displayed in Table 2. The 5-scale difference (ie, 25 dB difference) had the highest sensitivity (95.5%, 95% CI 87.5%-99.1%) in diagnosing SSNHL, while the 7-scale difference (ie, 35 dB difference) showed the highest specificity (90.5%, 95% CI 69.6%-98.8%).

Table 2. Diagnostic validity of the hearing scale differencea.

Hearing scale difference	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV[b], % (95% CI)	NPV[c], % (95% CI)
5 hearing scales (25 dB)	95.5 (87.5-99.1)	66.7 (43.0-85.4)	90.1 (80.7-95.9)	82.3 (56.6-96.2)
6 hearing scales (30 dB)	92.5 (83.4-97.5)	85.7 (63.7-97.0)	95.4 (87.1-99.0)	78.3 (56.3-92.5)
7 hearing scales (35 dB)	91.0 (81.5-96.6)	90.5 (69.6-98.8)	96.8 (89.0-99.6)	76.0 (54.9-90.6)

[a]Hearing scale difference = hearing scale of impaired ear – hearing scale of contralateral ear.

[b]PPV: positive predictive value.

[c]NPV: negative predictive value.

Discussion

Principal Findings

This study is the first to investigate the validity of a smartphone-based hearing screening app integrated with the novel HST for the assessment of SSNHL. This study confirmed that there is a strong correlation in hearing results between conventional pure-tone audiometry and the proposed smartphone-based Ear Scale app in a cohort of patients with possible SSNHL. The sensitivity of the hearing scale difference between the two ears measured by the Ear Scale app (95.5% for 5-hearing scale difference [25 dB] with a specificity of 66.7%, 92.5% for 6-hearing scale difference [30 dB] with a specificity of 85.7% , and 91.0% for 7-hearing scale difference [35 dB] with a specificity of 90.5%) in diagnosing SSNHL was high, suggesting that the smartphone-based approach can assist

in the evaluation of SSNHL, particularly in clinical settings where conventional pure-tone audiometry is not available.

Comparison With Prior Works

A previous study implemented the proposed Ear Scale app for hearing screening among a pediatric population and reported a strong correlation in the PTA between the app and conventional pure-tone audiometry in a soundproof booth as well as high accuracy in identifying school-aged children with hearing impairment [26]. Notably, Handzel et al [32] used a different smartphone-based app, the uHear hearing test app, for the initial assessment of unilateral SSNHL in 32 patients who had been diagnosed with SSNHL by standard audiometry. Table 3 illustrates the comparison between the gold standard approach (ie, pure-tone audiometry), the uHear app, and the proposed Ear Scale app in this study. The authors observed a sensitivity of 76% with the most stringent gold standard and of 94% with the least stringent criterion when they used the smartphone-based hearing screening tool for diagnosing SSNHL (Table 3) [32]. Our results were consistent with these findings, added to the literature by providing results in a larger sample size and better diagnostic validity and further broadening the population eligible for hearing screening using the hearing scale difference between the impaired ear and the contralateral ear as measured by the Ear Scale app. A significant strength of the proposed method for evaluating SSNHL is that instead of measuring the exact hearing thresholds, we used the hearing scale difference between the two ears to identify SSNHL. A major concern in measuring hearing status using these smartphone- or tablet-based tools is the ambient noise level, since they are not administered in a soundproof booth like conventional pure-tone audiometry. The presence of background noise can negatively affect hearing performance and lead to inaccurate results. This problem was minimized with our approach, as the hearing scale difference between the two ears was used and the influence of ambient noise was therefore canceled out. This unique feature indicates that the implementation of the Ear Scale app can be feasible in noisy environments, thereby broadening its applicability to settings such as urgent care clinics or emergency departments.

Table 3. Comparison of key characteristics among different approaches for identifying sudden sensorineural hearing loss.

Diagnostic approach	Author	Audiometric criteria of SSNHL[a]	Role	Sample size, n	Measurement unit	Sensitivity/specificity, %
Conventional pure-tone audiometry	Stachler et al [1]	A decrease in hearing of ≥30 dB, affecting at least 3 consecutive frequencies[b,c]	Gold standard	—[d]	dB HL[e]	—
uHear hearing test app	Handzel et al [31]	Hearing loss of at least 2 hearing grades across 3 or more consecutive frequencies[c,f]	Smartphone-based test	32	Hearing grade	76.0/91.0
Ear Scale app (current study)	Lin et al [10]	Hearing loss of at least 5 hearing scales difference[c]	Smartphone-based test	88	Hearing scale	95.5/66.7

[a]SSNHL: sudden sensorineural hearing loss.

[b]Definition according to the American Academy of Otolaryngology-Head and Neck Surgery guidelines [1].

[c]Hearing loss is defined as related to the opposite ear's thresholds.

[d]not available.

[e]dB HL: decibel hearing level.

[f]Hearing thresholds are grouped into 6 grades (American Speech-Language-Hearing Association 2012: normal 0-25 dB, mild 26-40 dB HL, moderate 41-55 dB HL, moderately severe 56-70 dB HL, severe 71-90 dB HL, profound >90 dB HL).

Clinical Implications

Several practice guidelines and reviews have suggested that patients with possible SSNHL undergo a comprehensive clinical workup upon arrival in the clinic, including thorough history taking, relevant physical examinations, and tuning fork tests to differentiate other types of hearing loss from SSNHL, identify nonidiopathic etiologies, and generate differential diagnoses [1,5,33,34]. Although these approaches are important and convenient, they can yield unreliable, even misleading, results [35,36]. Audiometric confirmation is still mandatory for definitively diagnosing SSNHL and should be performed on an emergent basis [1,5]. Conventional pure-tone audiometry remains the preferred method because it accurately distinguishes conductive hearing loss from those of sensorineural origins and establishes frequency-specific hearing thresholds, which are required components of frequently used audiometric criteria for SSNHL [1,5]. Initial audiometric outcomes also provide information essential for predicting prognoses and planning treatments [1]. Given their critical role in the management of SSNHL, audiometric assessments should be performed in accordance with the protocols proposed by the American Speech-Language-Hearing Association and the standards regarding the maximum allowable ambient noise and proper calibration [1,37,38]. In primary care practices (PCPs) or other busy clinical settings, such as urgent care and emergency departments, performing a standard battery of audiology tests can be challenging [13,39]. The high costs of equipment, limited space and time, noisy environments, and shortage of qualified personnel who are capable of conducting the screening and daily health assessments with audiometry are barriers to conventional pure-tone evaluations [13,39,40]. In a study that investigated the procedures performed by general practitioners working in PCPs, less than 20% of clinicians performed audiometry in their practices [41]. Since conventional pure-tone audiometry is mostly unavailable in PCP settings, innovative telehealth approaches have emerged that have been demonstrated as applicable and cost-effective for hearing assessments in

PCP-level settings [14,40,42,43]. The Ear Scale app proposed in this study has been shown to be feasible for hearing screenings in the pediatric population [26] and has been shown to have a good level of diagnostic accuracy for SSNHL. We believe that this novel tool, which incorporates the HST and smartphone-based technology, can serve as a point-of-care test for SSNHL at the PCP level because it is affordable, efficient, and requires minimal training to administer. The proposed procedure used in this study (Figure 2) could be the standardized approach when implementing the Ear Scale app in real-world settings. It could therefore assist health care providers in PCP or urgent care settings in making appropriate decisions regarding otolaryngology referrals, reduce the possibility of delayed management, and potentially improve the hearing recovery of patients with SSNHL.

Since the premorbid hearing status is generally unknown among people with possible SSNHL, hearing loss is usually defined on the basis of the difference between the two ears in the thresholds [1]. Based on our results, the PTA differences and hearing scale differences between the two ears are strongly correlated, with a correlation coefficient of .84. The diagnostic validity of three selected hearing scale difference cutoffs is reported in our study. Although all three cutoff values yielded satisfactory sensitivity, we preferred and recommended using the 5-hearing scale difference, as it had the lowest false-negative responses and can serve as the diagnostic standard for SSNHL. There is evidence that patients with untreated/unrecovered SSNHL have more tinnitus and balance problems as well as a poorer long-term quality of life [6,44]. These findings pose significant concerns regarding other negative health consequences associated with hearing loss, including falls [9,45], social isolation [46], depression [47], and incident dementia [10]. In the presence of other common sources of hearing loss, such as presbycusis, the impact of SSNHL is aggravated. In addition, misclassifying diseased cases as nondiseased cases may lead to delayed care among individuals with SSNHL, which is an important prognostic factor because it can be prevented [7,8]. Given that further hearing evaluations of SSNHL, which mainly include standard pure-tone audiometric assessments, are neither invasive nor harmful, minimizing the false-negative rate should be the goal of adequate tools when screening persons with possible SSNHL.

Limitations

Our results are limited by a high prevalence of SSNHL because the study cohort consisted of patients referred to tertiary academic medical centers. This factor could potentially bias our estimations of the positive and negative predictive values. Additionally, our screening approach may not be applicable for individuals with bilateral hearing loss. Individuals with conductive hearing loss were excluded from the study cohort, and testing such individuals may pose concerns of safety and feasibility. Despite these limitations, our study confirmed that hearing results can be compared between conventional pure-tone audiometry and the proposed Ear Scale app and that the app has sufficient diagnostic validity for SSNHL. To increase the generalizability and ensure the feasibility and safety of this smartphone-based hearing screening approach in PCP or urgent care settings, future studies with prospective designs and larger sample sizes, acceptability surveys among patients and clinicians, and health-economic analyses are needed. Furthermore, the proposed smartphone-based Ear Scale app creates new possibilities in the management of SSNHL at the PCP level since it can serve as a patient surveillance tool, enabling frequent monitoring and treatment adjustments [24,25]. Given the often limited insurance coverage of conventional pure-tone audiometric assessments, implementing a smartphone-based hearing screening approach is a crucial step toward the decentralization of hearing care in the PCP setting, increased accessibility of timely management, and, ultimately, better hearing prognoses in patients with SSNHL.

Conclusions

This study demonstrated that the hearing results measured by conventional pure-tone audiometry and the proposed smartphone-based Ear Scale app are strongly correlated among patients with possible SSNHL. Our results showed that the hearing scale difference between the two ears, as measured by the Ear Scale app, has a satisfactory level of validity in detecting SSNHL. The results also suggested that this smartphone-based approach may effectively assist the evaluation of SSNHL in clinical settings where conventional pure-tone audiometry is not available.

Acknowledgments

The authors would like to thank Shang-Liang Wu, Hsia-Wei Tseng, Hsin-Yi Huang, the Big Data Center of Taipei Veterans General Hospital, and the Biostatistics Task Force of Taipei Veterans General Hospital for their assistance during the study. This study was based in part on data from the Big Data Center at Taipei Veterans General Hospital. The interpretations and conclusions contained herein do not represent the position of the Taipei Veterans General Hospital. This study was supported by grants from Taipei Veterans General Hospital (V108C-178, V108C-145, V109C-157, 108VACS-003), the Ministry of Health and Welfare (MOST 108-2221-E-075-001, MOHW109-TDU-B-211-114001, MOST 109-2221-E-075-001, MOHW109-TDU-B-211-134019) and the Ministry of Science and Technology (MOST108-2221-E-075-001, MOST109-2320-B-075-006, MOST109-2218-E-010-004). The funders did not play any roles in the study design, data collection or analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

Both WHL and YFC (email: yfcheng2@vghtpe.gov.tw) are corresponding authors for this paper and contributed equally to this work. WHL and YFC were responsible for the acquisition of the data. HYHL, YCC, YHL, and HLC were responsible for the analysis and interpretation of the data. HYHL and YCC drafted of the manuscript. HYHL and YCC performed the statistical analysis. WHL and YFC obtained funding, and YCC, YHL, FL, and YFC were responsible for administrative, technical, or material support. YFC and WHL supervised the study. All authors were responsible for study concept and design, critical revision of the manuscript for important intellectual content, and approval of the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sound volume in decibels for each stimulation level examined in the Hearing Scale Test and for different frequencies.
[DOCX File , 14 KB - mhealth_v8i11e23047_app1.docx]

References

1. Stachler RJ, Chandrasekhar SS, Archer SM, Rosenfeld RM, Schwartz SR, Barrs DM, et al. Clinical practice guideline: sudden hearing loss. Otolaryngol Head Neck Surg 2012 Mar;146(3 Suppl):S1-S35. [doi: 10.1177/0194599812436449] [Medline: 22383545]
2. Alexander TH, Harris JP. Incidence of sudden sensorineural hearing loss. Otol Neurotol 2013 Dec;34(9):1586-1589. [doi: 10.1097/MAO.0000000000000222] [Medline: 24232060]
3. Byl FM. Seventy-six cases of presumed sudden hearing loss occurring in 1973: prognosis and incidence. Laryngoscope 1977 May;87(5 Pt 1):817-825. [doi: 10.1002/lary.5540870515] [Medline: 850455]
4. Kuo C, Chung C, Wang C, Chien W, Chen H. Increased incidence in hospitalised patients with sudden sensorineural hearing loss: a 14-year nationwide population-based study. Int J Audiol 2019 Nov;58(11):769-773. [doi: 10.1080/14992027.2019.1627009] [Medline: 31195853]
5. Kuhn M, Heman-Ackah SE, Shaikh JA, Roehm PC. Sudden sensorineural hearing loss: a review of diagnosis, treatment, and prognosis. Trends Amplif 2011 Sep;15(3):91-105 [FREE Full text] [doi: 10.1177/1084713811408349] [Medline: 21606048]
6. Härkönen K, Kivekäs I, Rautiainen M, Kotti V, Vasama J. Quality of life and hearing eight years after sudden sensorineural hearing loss. Laryngoscope 2017 Apr;127(4):927-931. [doi: 10.1002/lary.26133] [Medline: 27328455]
7. Edizer DT, Çelebi O, Hamit B, Baki A, Yiğit O. Recovery of idiopathic sudden sensorineural hearing loss. J Int Adv Otol 2015 Aug;11(2):122-126 [FREE Full text] [doi: 10.5152/iao.2015.1227] [Medline: 26381001]
8. Cvorović L, Deric D, Probst R, Hegemann S. Prognostic model for predicting hearing recovery in idiopathic sudden sensorineural hearing loss. Otol Neurotol 2008 Jun;29(4):464-469. [doi: 10.1097/MAO.0b013e31816fdcb4] [Medline: 18434930]
9. Deal JA, Richey Sharrett A, Bandeen-Roche K, Kritchevsky SB, Pompeii LA, Gwen Windham B, et al. Hearing impairment and physical function and falls in the atherosclerosis risk in communities hearing pilot study. J Am Geriatr Soc 2016 Apr;64(4):906-908 [FREE Full text] [doi: 10.1111/jgs.14075] [Medline: 27100600]
10. Lin FR, Metter EJ, O'Brien RJ, Resnick SM, Zonderman AB, Ferrucci L. Hearing loss and incident dementia. Arch Neurol 2011 Feb;68(2):214-220 [FREE Full text] [doi: 10.1001/archneurol.2010.362] [Medline: 21320988]
11. Deal JA, Sharrett AR, Albert MS, Coresh J, Mosley TH, Knopman D, et al. Hearing impairment and cognitive decline: a pilot study conducted within the atherosclerosis risk in communities neurocognitive study. Am J Epidemiol 2015 May 01;181(9):680-690 [FREE Full text] [doi: 10.1093/aje/kwu333] [Medline: 25841870]
12. Reed NS, Altan A, Deal JA, Yeh C, Kravetz AD, Wallhagen M, et al. Trends in health care costs and utilization associated with untreated hearing loss over 10 years. JAMA Otolaryngol Head Neck Surg 2019 Jan 01;145(1):27. [doi: 10.1001/jamaoto.2018.2875]
13. Louw C, Swanepoel DW, Eikelboom RH. Self-reported hearing loss and pure tone audiometry for screening in primary health care clinics. J Prim Care Community Health 2018;9:2150132718803156. [doi: 10.1177/2150132718803156] [Medline: 30278815]
14. Clark JL, Swanepoel DW. Technology for hearing loss – as we know it, and as we dream it. Disabil Rehabil Assistive Technol 2014 Apr 08;9(5):408-413. [doi: 10.3109/17483107.2014.905642]
15. Ratanjee-Vanmali H, Swanepoel DW, Laplante-Lévesque A. Patient uptake, experience, and satisfaction using web-based and face-to-face hearing health services: process evaluation study. J Med Internet Res 2020 Mar 20;22(3):e15875 [FREE Full text] [doi: 10.2196/15875] [Medline: 32196459]
16. Masalski M, Kręcicki T. Self-test web-based pure-tone audiometry: validity evaluation and measurement error analysis. J Med Internet Res 2013;15(4):e71 [FREE Full text] [doi: 10.2196/jmir.2222] [Medline: 23583917]

17. Liao W, Young S, Lien C, Wang S. An audiometer to monitor progressive hearing change in school-aged children. J Med Screen 2011;18(1):8-11. [doi: 10.1258/jms.2011.010120] [Medline: 21536810]
18. Honeth L, Bexelius C, Eriksson M, Sandin S, Litton J, Rosenhall U, et al. An internet-based hearing test for simple audiometry in nonclinical settings: preliminary validation and proof of principle. Otol Neurotol 2010 Jul;31(5):708-714. [doi: 10.1097/MAO.0b013e3181de467a] [Medline: 20458255]
19. Yimtae K, Israsena P, Thanawirattananit P, Seesutas S, Saibua S, Kasemsiri P, et al. A tablet-based mobile hearing screening system for preschoolers: design and validation study. JMIR Mhealth Uhealth 2018 Oct 23;6(10):e186 [FREE Full text] [doi: 10.2196/mhealth.9560] [Medline: 30355558]
20. Masalski M, Grysiński T, Kręcicki T. Hearing tests based on biologically calibrated mobile devices: comparison with pure-tone audiometry. JMIR Mhealth Uhealth 2018 Jan 10;6(1):e10 [FREE Full text] [doi: 10.2196/mhealth.7800] [Medline: 29321124]
21. Sandström J, Swanepoel DW, Carel MH, Laurent C. Smartphone threshold audiometry in underserved primary health-care contexts. Int J Audiol 2016;55(4):232-238. [doi: 10.3109/14992027.2015.1124294] [Medline: 26795898]
22. Masalski M, Kipiński L, Grysiński T, Kręcicki T. Hearing tests on mobile devices: evaluation of the reference sound level by means of biological calibration. J Med Internet Res 2016 May 30;18(5):e130 [FREE Full text] [doi: 10.2196/jmir.4987] [Medline: 27241793]
23. Bright T, Pallawela D. Validated smartphone-based apps for ear and hearing assessments: a review. JMIR Rehabil Assist Technol 2016 Dec 23;3(2):e13 [FREE Full text] [doi: 10.2196/rehab.6074] [Medline: 28582261]
24. Liao W, Lien C, Young S. The Hearing Scale Test for hearing screening of school-age children. Int J Pediatr Otorhinolaryngol 2010 Jul;74(7):760-764. [doi: 10.1016/j.ijporl.2010.03.033] [Medline: 20394995]
25. Liao W, Young S, Tang S. A novel method for quick hearing assessment of children. ICEIE 2010 Sep;1:1-37. [doi: 10.1109/iceie.2010.5559833]
26. Chu Y, Cheng Y, Lai Y, Tsao Y, Tu T, Young ST, et al. A mobile phone-based approach for hearing screening of school-age children: cross-sectional validation study. JMIR Mhealth Uhealth 2019 Apr 01;7(4):e12033 [FREE Full text] [doi: 10.2196/12033] [Medline: 30932870]
27. Cheng Y, Chu Y, Tu T, Shiao A, Wu S, Liao W. Modified Siegel's criteria for sudden sensorineural hearing loss: reporting recovery outcomes with matched pretreatment hearing grades. J Chin Med Assoc 2018 Nov;81(11):1008-1012 [FREE Full text] [doi: 10.1016/j.jcma.2018.03.012] [Medline: 29929833]
28. Foulad A, Bui P, Djalilian H. Automated audiometry using apple iOS-based application technology. Otolaryngol Head Neck Surg 2013 Nov;149(5):700-706. [doi: 10.1177/0194599813501461] [Medline: 23963611]
29. Ho C, Li P, Young S. Reference equivalent threshold sound pressure levels for Apple EarPods. J Acoust Soc Am 2017 Feb;141(2):EL115-EL119. [doi: 10.1121/1.4976110] [Medline: 28253642]
30. Gardner WG, Martin KD. HRTF measurements of a KEMAR. J Acoust Soc Am 1995 Jun;97(6):3907-3908. [doi: 10.1121/1.412407]
31. Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am 1971 Feb;49(2):467-477. [Medline: 5541744]
32. Handzel O, Ben-Ari O, Damian D, Priel MM, Cohen J, Himmelfarb M. Smartphone-based hearing test as an aid in the initial evaluation of unilateral sudden sensorineural hearing loss. Audiol Neurootol 2013;18(4):201-207. [doi: 10.1159/000349913] [Medline: 23689282]
33. Jensen EAH, Harmon ED, Smith W. Early identification of idiopathic sudden sensorineural hearing loss. Nurse Pract 2017 Sep 21;42(9):10-16. [doi: 10.1097/01.NPR.0000521990.15440.ee] [Medline: 28787350]
34. Burgess LP, Frankel SF, Lepore ML, Yim DW. Tuning fork screening for sudden hearing loss. Mil Med 1988 Sep;153(9):456-458. [Medline: 3141836]
35. Stankiewicz JA, Mowry HJ. Clinical accuracy of tuning fork tests. Laryngoscope 1979 Dec;89(12):1956-1963. [doi: 10.1288/00005537-197912000-00009] [Medline: 513917]
36. Browning GG, Swan IR, Chew KK. Clinical role of informal tests of hearing. J Laryngol Otol 1989 Jan;103(1):7-11. [doi: 10.1017/s0022215100107923] [Medline: 2646384]
37. American Speech-Language-Hearing Association. Preferred Practice Patterns for the professions of speech-language pathology and audiology. ASHA Suppl 1993 Mar;35(3 Suppl 11):1-102. [Medline: 8097652]
38. Frank T. ANSI update: maximum permissible ambient noise levels for audiometric test rooms. Am J Audiol 2000 Jun;9(1):3-8. [doi: 10.1044/1059-0889(2000/003)]
39. Bettger JP, Dolor RJ, Witsell DL, Dubno JR, Pieper CF, Walker AR, et al. Comparative implementation-effectiveness of three strategies to perform hearing screening among older adults in primary care clinics: study design and protocol. BMC Geriatr 2020 May 11;20(1):170 [FREE Full text] [doi: 10.1186/s12877-020-01576-x] [Medline: 32393184]
40. Louw C, Swanepoel DW, Eikelboom RH, Myburgh HC. Smartphone-based hearing screening at primary health care clinics. Ear Hear 2017;38(2):e93-e100. [doi: 10.1097/AUD.0000000000000378] [Medline: 27764002]
41. Flaegel K, Brandt B, Goetz K, Steinhaeuser J. Which procedures are performed by general internists practicing primary care in Germany? A cross-sectional study. BMC Fam Pract 2020 Apr 29;21(1). [doi: 10.1186/s12875-020-01136-7]

42. Swanepoel DW, Myburgh HC, Howe DM, Mahomed F, Eikelboom RH. Smartphone hearing screening with integrated quality control and data management. Int J Audiol 2014 Dec;53(12):841-849. [doi: 10.3109/14992027.2014.920965] [Medline: 24998412]
43. Yousuf Hussein S, Wet Swanepoel D, Biagio de Jager L, Myburgh HC, Eikelboom RH, Hugo J. Smartphone hearing screening in mHealth assisted community-based primary care. J Telemed Telecare 2016 Oct;22(7):405-412. [doi: 10.1177/1357633X15610721] [Medline: 26468215]
44. Chiossoine-Kerdel JA, Baguley DM, Stoddart RL, Moffat DA. An investigation of the audiologic handicap associated with unilateral sudden sensorineural hearing loss. Am J Otol 2000 Sep;21(5):645-651. [Medline: 10993452]
45. Lin FR, Ferrucci L. Hearing loss and falls among older adults in the United States. Arch Intern Med 2012 Feb 27;172(4):369-371 [FREE Full text] [doi: 10.1001/archinternmed.2011.728] [Medline: 22371929]
46. Shukla A, Harper M, Pedersen E, Goman A, Suen JJ, Price C, et al. Hearing loss, loneliness, and social isolation: a systematic review. Otolaryngol Head Neck Surg 2020 May;162(5):622-633. [doi: 10.1177/0194599820910377] [Medline: 32151193]
47. Shukla A, Reed N, Armstrong NM, Lin FR, Deal JA, Goman AM. Hearing loss, hearing aid use and depressive symptoms in older adults - findings from the Atherosclerosis Risk in Communities Neurocognitive Study (ARIC-NCS). J Gerontol B Psychol Sci Soc Sci 2019 Oct 19. [doi: 10.1093/geronb/gbz128] [Medline: 31628485]

Abbreviations

dB: decibel
HL: hearing level
HST: Hearing Scale Test
KEMAR: Knowles Electronics Manikin for Acoustic Research
PCP: primary care practice
PTA: pure-tone average
SSNHL: sudden sensorineural hearing loss

Edited by G Eysenbach; submitted 30.07.20; peer-reviewed by CS Wu, O Jones, M Masalski; comments to author 17.09.20; revised version received 10.10.20; accepted 21.10.20; published 11.11.20.

Please cite as:
Lin HYH, Chu YC, Lai YH, Cheng HL, Lai F, Cheng YF, Liao WH
A Smartphone-Based Approach to Screening for Sudden Sensorineural Hearing Loss: Cross-Sectional Validity Study
JMIR Mhealth Uhealth 2020;8(11):e23047
URL: http://mhealth.jmir.org/2020/11/e23047/
doi:10.2196/23047
PMID:33174845

Original Paper

Effect of a Text Messaging–Based Educational Intervention on Cesarean Section Rates Among Pregnant Women in China: Quasirandomized Controlled Trial

Yanfang Su[1*], SCD; Jesse Heitner[2*], SCD; Changzheng Yuan[3,4*], SCD; Yafei Si[5], MA; Dan Wang[6], MA; Zhiying Zhou[7], MA; Zhongliang Zhou[6], PhD

[1]School of Medicine, University of Washington, Seattle, WA, United States

[2]Aceso Global, Washington, DC, United States

[3]The Children's Hospital and School of Public Health, Zhejiang University School of Medicine, Hangzhou, China

[4]Nutrition Department, Harvard School of Public Health, Boston, MA, United States

[5]School of Risk & Actuarial Studies and Centre of Excellence in Population Ageing Research (CEPAR), University of New South Wales, Sydney, Australia

[6]School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an, China

[7]School of Public Health, Xi'an Jiaotong University Health Science Center, Xi'an, China

[*]these authors contributed equally

Corresponding Author:
Zhongliang Zhou, PhD
School of Public Policy and Administration
Xi'an Jiaotong University
28# Xianning West Road
Xi'an
China
Phone: 86 18291498261
Email: zzliang1981@xjtu.edu.cn

Abstract

Background: Consensus exists that appropriate regional cesarean rates should not exceed 15% of births, but China's cesarean rate exceeds 50% in some areas, prompting numerous calls for its reduction. At present, China's 2016 two-child policy has heightened the implications of national cesarean section trends.

Objective: This study leveraged pervasive cellular phone access amongst Chinese citizens to test the effect of a low-cost and scalable prenatal advice program on cesarean section rates.

Methods: Participants were pregnant women presenting for antenatal care at a clinic in Xi'an, China. Assignment was quasirandomized and utilized factorial assignment based on the expecting mother's birthday. Participants were assigned to one of the following four groups, with each receiving a different set of messages: (1) a comparison group that received only a few "basic" messages, (2) a group receiving messages primarily regarding care seeking, (3) a group receiving messages primarily regarding good home prenatal practices, and (4) a group receiving text messages of all groups. Messages were delivered throughout pregnancy and were tailored to each woman's gestational week. The main outcome was the rates of cesarean delivery reported in the intervention arms. Data analysts were blinded to treatment assignment.

Results: In total, 2115 women completed the trial and corresponding follow-up surveys. In the unadjusted analysis, the group receiving all texts was associated with an odds ratio of 0.77 (P=.06), though neither the care seeking nor good home prenatal practice set yielded a relevant impact. Adjusting for potentially confounding covariates showed that the group with all texts sent together was associated with an odds ratio of 0.67 (P=.01). Notably, previous cesarean section evoked an odds ratio of 11.78 (P<.001), highlighting that having a cesarean section predicts future cesarean section in a subsequent pregnancy.

Conclusions: Sending pregnant women in rural China short informational messages with integrated advice regarding both care-seeking and good home prenatal practices appears to reduce women's likelihood of undergoing cesarean section. Reducing clear medical indications for cesarean section seems to be the strongest potential pathway of the effect. Cesarean section based

on only maternal request did not seem to occur regularly in our study population. Preventing unnecessary cesarean section at present may have a long-term impact on future cesarean section rates.

Trial Registration: ClinicalTrials.gov NCT02037087; https://clinicaltrials.gov/ct2/show/NCT02037087.

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2015-011016

(JMIR Mhealth Uhealth 2020;8(11):e19953) doi:10.2196/19953

KEYWORDS

cesarean section; short message service; SMS text messaging; quasirandomized controlled trial; mobile health

Introduction

The global health care community has estimated that regional cesarean section (CS) rates should not exceed 10% to 15% [1,2]. However, in the Global Survey by the World Health Organization (WHO), CS in China was estimated to involve 46.2% of all deliveries, which is the highest rate for any country in the survey [3]. In a recent multicenter survey of 39 hospitals across mainland China, the overall CS rate was 54.90% [4]. It is estimated that between 1990 and 2014, China had an average annual rate of increase in CS of about 10%, and misconceptions of pain, genital modification, safety, and cultural fortune were reported as leading factors [5]. Since at least 2008, this increase was most visible in the rural counties of China [6]. Though rural counties have clearly lower rates of CS than general city or "supercity" areas, as of 2014, the average rural county CS rate was slightly over 30%, and it is rising [6].

When warranted, CS is a vital intervention. Many studies have confirmed that it has a strong protective effect on perinatal mortality when conditions, such as breech presentation, placenta previa, and uterine rupture, are encountered [3,7,8]. A study of 66,226 deliveries in Shanghai found that compared with vaginal delivery, CS was associated with a reduction in antepartum stillbirth, bone trauma, intracranial hemorrhage, and neonatal hypoxemic encephalopathy [9]. Moreover, studies suggest no evidence of a difference in maternal mortality between *planned* vaginal and *planned* cesarean delivery [10]. In the Global Survey by the WHO, the maternal mortality risk for antepartum CS without an indication could not be estimated because there were no maternal deaths in the group [3]. However, CS does come with serious risks [8]. Using data from the Global Survey by the WHO, although researchers found no association with maternal *mortality*, CS without medical indications had strong associations with severe maternal *morbidity*. On putting death and several severe morbidities, namely admission to the intensive care unit, blood transfusion, and hysterectomy, into one "Severe Maternal Outcomes" index, the authors found that elective antepartum CS had an adjusted odds ratio (OR) of 5.93 (CI 3.88-9.05) for qualifying for the index and elective intrapartum CS had an adjusted OR of 14.29 (CI 10.91-18.72) (both $P<.05$) [11]. A large cohort study in Australia found that mothers delivering via CS were more likely to be readmitted to the hospital within 8 weeks of birth [12].

CS is also associated with problems after delivery. A recent study found that women delivering via CS had roughly twice the odds of persistent pain 1 year after delivery [13]. The association between CS and reduced future fertility has been demonstrated in numerous studies [14]. In a new pregnancy, a *prior* CS may cause an increased risk of fetal wastage and may be linked to unexplained stillbirth [10,14]. Further, there is a strong body of evidence on impaired uterine function following cesarean delivery. In subsequent pregnancies, CS poses a risk of uterine scar dehiscence, and, in some cases, uterine rupture [14]. As of January 2016, China altered its one-child policy to a two-child policy, suddenly making the effects of CS on subsequent pregnancies a tremendously more important consideration.

Given the risks and benefits, the WHO has concluded that "cesarean section should ideally only be undertaken when medically necessary" [2]. Yet, Lumbiganon et al estimated that 11.7% of all deliveries in China during their study period involved CS with no medical indications [3]. Combining the 24 countries in the Global Survey by the WHO, it was estimated that 63% of all CS procedures without medical indications were performed in China [11]. An important driver of this pattern is that women with no indications necessitating CS frequently request this procedure. Cesarean delivery on maternal request accounted for 15.53% of all deliveries and 28.43% of all cesarean deliveries in a multicenter survey [4]. This national estimate confirms what at least 11 other smaller and qualitative studies [15] and at least one regional estimate [16] have suggested (women's preferential choices are part of the rise in China's CS rates). Influencing this demand may be an important strategy for reducing medically unnecessary CS.

There is also evidence that part of the "demand" for CS is supplier induced [17]. CS brings in approximately double the hospital revenue per birth as compared with vaginal delivery [16,18], and the power imbalance between patients and providers may mask the true decision making [14,15]. In a recent study in Shanghai, of 599 women interviewed in their third trimester, 17.0% reported preferring cesarean delivery. Yet, among women completing the study, 58.1% underwent CS. Of those, only 50.0% had clinically accepted indications for CS [19]. Therefore, educating and empowering women to refute inappropriate doctor recommendations for CS may be as important of a pathway for reducing CS as changing women's underlying preferences.

There is agreement that the rate of CS in China, whether supply or demand-side driven, is excessive, and experts are calling for strategies for its reduction [1,4,17]. Mobile health (mHealth) generally has already shown relevant effects in several intervention areas. However, evidence for or against its efficacy for maternal and child health is scarce [20,21], and larger scale evaluations of its possible effects are warranted [22-27]. This study comprises a portion of the Evaluation for mHealth

Interventions' Newborn Health Project. The Newborn Health Project has several aims, and its primary (trial registered) metric of success is a newborn's appropriate weight for gestational age [28,29]. This study investigated whether the project was successful in the secondary goal of lowering the rate of CS in the intervention arms.

Methods

The Newborn Health Project offers expectant mothers in the rural district of Gaoling in Xi'an, China, a package of free, short, informational messages regarding pregnancy and childbirth via cell phones. The full protocol of the Newborn Health Project has been published elsewhere [29], including the process and rationale for selecting the study site. The study utilizes factorial quasirandomization at the individual level to assign women to receive one of four groups of text messages and then compares outcomes between the four groups. Participating women were blinded to assignment. All data analyses were performed with blinding to treatment assignment. The trial has been registered at ClinicalTrials.gov (NCT02037087).

The four study arms were as follows: (1) good home prenatal practice messages (home practices), which included advice on nutrition, exercise, self-awareness of depression, breastfeeding, etc; (2) care-seeking messages (care seeking), which included information about government-subsidized programs, warning signs of potential problems, and the importance of care seeking during illness; (3) both types of messaging (all texts); and (4) a very limited (25 in total) set of "basic" messages about pregnancy (acting as a comparison group). Women in the other intervention arms also received all of the basic messages. It was decided that the comparison group should receive at least some regular informational "placebo" messages to make the participants in this group feel like they received a service and were part of the program. Ethically, it also ensured that all enrollees received the most basic pregnancy information, which is the informational equivalent of "basic care." These basic messages primarily included updates on fetal development, as well as reminders for prenatal visits and promotion of certified skilled attendance of labor. Group comparisons of treatment arms elicited the effect of (assignment to) receiving the *content* in the intervention messages in addition to the basic ones, and estimated this effect separated out from any effect associated with being included in an informational messaging study. The number of messages by topic and study arm is presented in Multimedia Appendix 1.

The four treatment arms received differing sets of messages relevant to labor and delivery that could potentially impact a woman's choice of delivery mode. Of the "basic" (comparison) group's 25 messages, none were relevant to delivery. The "care-seeking" group was sent seven relevant messages, generally focusing on describing proper indications for CS and cautioning that CS and anesthesia make birth less painful but come with their other risks. The "home practices" group was sent 15 delivery-relevant messages, generally focusing on inspiring confidence in vaginal delivery and discussing nonanesthetic ways to cope with pain during delivery. The "all texts" group received 22 relevant messages composed of those sent to the other treatment arms. The sent messages relevant to labor and delivery are presented in Multimedia Appendix 2. Though presented in English there, all sent messages were in Mandarin.

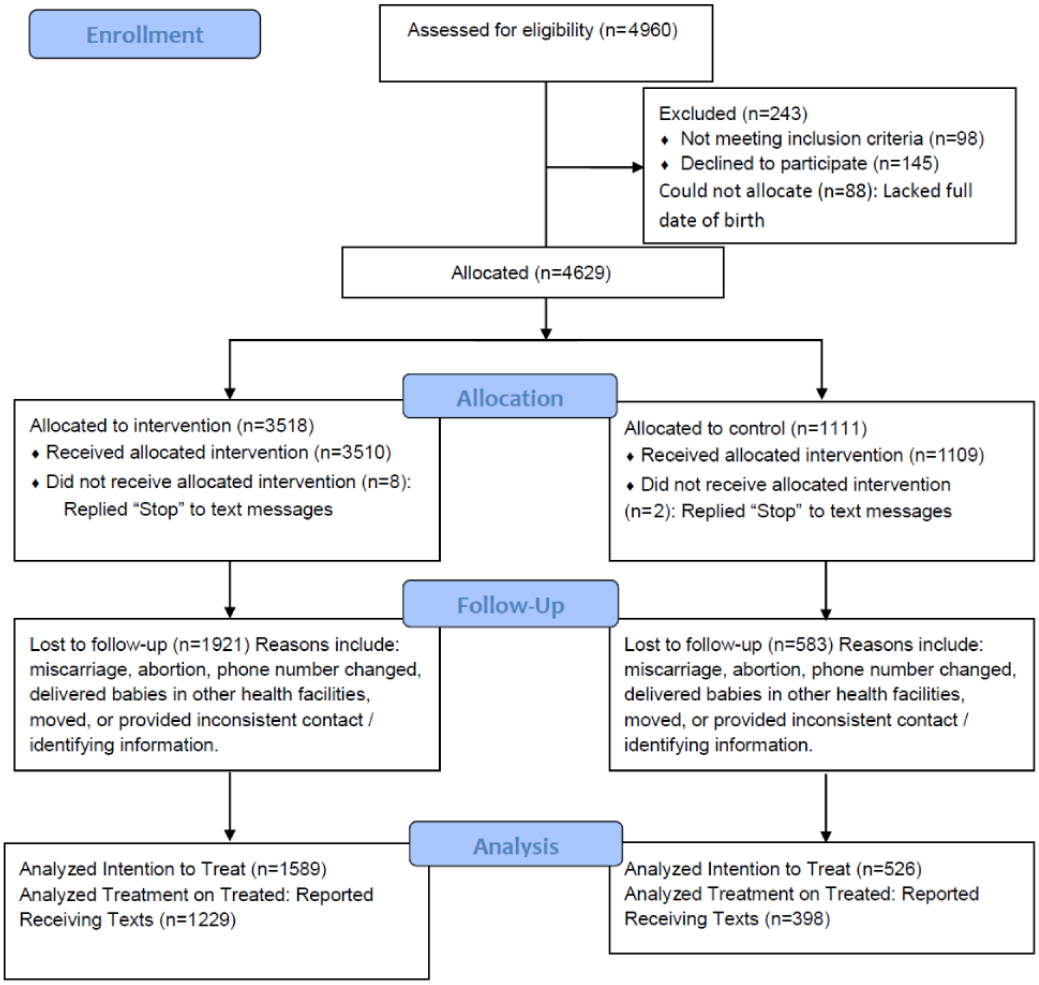
This study was approved by the Ethics Committee of the School of Medicine at Xi'an Jiaotong University on January 18, 2013, and an updated version was approved on May 2016 (approval number: 2016-392). All women attending their first visit for antenatal care at Gaoling district's local maternal and child health center (MCHC) during the study period were invited to participate if they were aged 18 to 45 years old and had access to a cellular phone within their household. All participants signed an informed consent form. Pilot testing with 140 subjects occurred between September and October 2013, and survey questionnaires were finalized after incorporating feedback from the testing. The participants included in our pilot were excluded in our data analysis and causal inferences. Enrollment and data collection for the study were performed between November 2013 and February 2016. Enrollment stopped before our power calculation–based sample size goals were reached in December 2015, when Gaoling MCHC stopped sending program texts after decided that future patient communications should preferably be sent over WeChat than cellular SMS text messaging. It is extremely noteworthy that the contents of our SMS intervention could easily be sent via WeChat, but in December 2015, the trial was not set up to do so.

At recruitment, prior to treatment assignment, a baseline survey was conducted by a health worker at the local MCHC during the first antenatal visit. We collected demographic data, self-reported health data, and data relating to each enrollee's thoughts and perceptions regarding health during pregnancy and childbirth. The follow-up survey was conducted by a health worker at the newborn's home around 1 week after delivery. The follow-up survey collected information on knowledge, psychological and behavioral changes, and pregnancy-related maternal and neonatal measures. A final survey was conducted by phone to assess postpartum depression around 1 month after delivery. Additionally, the final survey asked whether the enrollee had successfully received our messages.

Results

Our program enrolled a total of 4629 women (Figure 1). In our prior publication and Multimedia Appendix 3 [28], we present summary statistics and balance checks for all 55 measured baseline covariates. As previously reported via these balance checks, we inferred that our quasirandomization was effective in assigning treatment orthogonally to relevant observable covariates.

Figure 1. CONSORT flow diagram: enrollment, allocation, follow-up, and analysis.



In total, 2115 women completed a postdelivery follow-up survey, which could be linked to the baseline survey (Figure 1). Among these, 526 (24.9%) were in the "basic" group, 518 (24.5%) in the care-seeking group, 497 (23.5%) in the home practices group, and 574 (27.1%) in the all texts group. A chi-square test failed to reject equal loss to follow-up at P=.44. A balance check on all measured baseline characteristics was performed restricted to the final 2115 women. Only one test rejected balance at P<.05, and a further three rejected balance at .05<P<.10. We therefore inferred that among the women completing our study, our quasirandomization was effective in assigning treatment orthogonally to observable covariates. This set of balance checks is presented in Multimedia Appendix 4.

Loss to follow-up occurred via multiple pathways. First, only women who delivered their children, including stillbirths, were available for analysis regarding mode of delivery. This analysis was inapplicable for any woman who miscarried or underwent an abortion after enrollment. About 15% of clinically recognizable pregnancies end in spontaneous miscarriage within the first trimester [30]. As of 2010, China's induced abortion rate was hovering steadily around 19.5% [31]. Rates of miscarriage and abortion were not collected as part of our trial, but likely account for important shares of our loss to follow-up. Further, if women moved or otherwise gave birth outside of our study catchment area or if the village health attendant was otherwise unaware of a birth or unable to reach a new mother at her home, the pregnant woman was lost to follow-up. Finally, it was sometimes impossible to match follow-up surveys to baseline surveys. This happened when women provided sufficiently different identifying information in the baseline and follow-up surveys or when such information was sufficiently misrecorded.

The unadjusted rates of cesarean delivery by treatment assignment are presented in Table 1. The largest difference was between the "basic" group and the "all texts" group, whose rate was lower by 5.2 percentage points.

Table 1. Cesarean section numbers, rates, and odds ratios by treatment assignment, among women who reported delivery mode.

Group	Women who completed follow-up survey, n	Women who reported delivery mode, n	Vaginal deliveries, n	Cesarean sections, n (%)	Cesarean, odds ratio
Basic	526	522	372	150 (28.7)	Base case
Care Seeking	518	515	379	136 (26.4)	0.88
Home Practices	497	494	>369	>125 (25.3)	0.84
All Texts	574	>570	>436	>134 (23.5)	>0.77
Total	2115	2101	1556	545 (25.9)	N/A[a]

[a]N/A: not applicable.

Only 2101 of 2115 women reported their mode of delivery (Figure 1). Some missingness is to be expected in any large-scale survey, and no variable measured had all 2115 responses. It has been shown that a process called "multiple imputation" using expectation maximization will generally outperform the most common general techniques of handling missing data [32]. Thus, multiple imputation was performed in R (R Foundation for Statistical Computing) using the Amelia package. This process imputed 16 data sets that had "complete" data on all variables of interest, allowing all 2115 observations to be used in regression analysis. All regression analyses were run once per imputed data set, and the results were combined using the Rubin technique for combining quantities of interest [32].

Four regression models were run to explore the impact of treatment on CS rates. Model 1 is an unadjusted logistic regression of the (log) odds of having CS on indicators for assignment to each intervention arm, with the "basic" arm omitted as the base case. Results are presented in Table 2.

In Model 1, neither the "care-seeking" group nor the "home practices" group alone was associated with a statistically significant reduction in the odds of undergoing CS (P=.37 and P=.21, respectively). In combination, the all texts group was associated with an OR of 0.77, but the P value was .06.

A second model, also presented in Table 2, included all baseline covariates found to be unbalanced at P<.10 for either the full sample or for the 2115 subjects who completed the study. Adding these unbalanced covariates had negligible effect on regression results. The P value on assignment to the all texts group reached .048, but accounting for multiple comparisons left this finding not statistically significant. One noteworthy finding is that having a previous miscarriage was strongly predictive of undergoing CS at our study site, with an associated OR of 1.37 (95% CI 1.12-1.68). Whether the mechanism of this association is physiological or psychological may be an area for future study.

As a further robustness check, a third logistic regression model (Model 3) was run, which contained all unbalanced baseline covariates as well as an array of baseline general health, maternal, socioeconomic, and health psychology covariates that might influence or predict birth via CS. These results are also presented in Table 2. The health psychology covariates (not shown) attempted to account for the major constructs of the most widely cited theories of health behavior, namely the health belief model, social cognitive theory, theory of planned behavior, theory of reasoned action, and trans-theoretical model [33-35].

Table 2. Logistic regression analysis of cesarean birth according to treatment assignment.

Model	Odds ratio	95% CI	*P* value
Model 1: Unadjusted logistic regression			
Treatment assignment			
Basic only	Base case	N/A[a]	N/A
Care seeking	0.88	0.67-1.16	.37
Good home practices	0.84	0.63-1.11	.21
All texts	0.77	0.59-1.01	.06
Model 2[b]: Adjusted logistic regression for imbalance			
Treatment assignment			
Basic only	Base case	N/A	N/A
Care seeking	0.87	0.66-1.14	.32
Good home practices	0.82	0.62-1.08	.16
All texts	0.76	0.58-1.00	.048
Model 3[c]: Adjusted logistic regression for all			
Treatment assignment			
Basic only	Base case	N/A	N/A
Care seeking	0.75	0.54-1.04	.08
Good home practices	0.77	0.56-1.07	.12
All texts	0.67	0.49-0.92	.01

[a]N/A: not applicable.

[b]Model 2 adjusted for all unbalanced baseline covariates.

[c]Model 3 adjusted for all unbalanced baseline covariates as well as an array of baseline general health, maternal, socioeconomic, and health psychology covariates (Multimedia Appendix 4).

These additional covariates noticeably strengthened the measured effect of assignment to each treatment arm. The OR for the care-seeking group fell from 0.87 to 0.75, the OR for the home practice group fell from 0.82 to 0.77, and the OR for the all texts group fell from 0.76 to 0.67. The *P* values on assignment to the first two groups remained above traditional significance (*P*=.08 and *P*=.12 respectively), but that of assignment to the all text group dropped to *P*=.01, which remained statistically significant under the Bonferroni method of correcting for multiple testing of three tests in the same regression.

Most of the magnitude of these changes can be accounted for by the addition of two variable sets (analysis not shown), including whether the woman previously had CS (OR 11.8, 95% CI 7.2-19.3) and indicators for her stated preferred mode of delivery during enrollment. Preferring CS over vaginal birth and feeling unsure are both associated with about twice the odds of CS (OR 2.0, 95% CI 1.3-3.2 and OR 2.0, 95% CI 1.4-2.9, respectively). Intriguingly, this equality implies that feeling "unsure" carries the same effect as actually preferring CS, perhaps implicating a strong supply-side nudge toward CS for uncertain women. An enormously pressing implication of Model 3 is that if a woman has CS in one pregnancy, she is vastly more likely to have one in any subsequent pregnancy, raising the possibility that preventing unnecessary CS now may have a direct and long-term impact on future CS rates.

As mentioned, the postdelivery survey also inquired whether the enrollees had actually received text messages from the Newborn Health Project during pregnancy. Surprisingly, of the 2115 participants, only 1627 (76.93%) answered "Yes," 459 (21.70%) reported not receiving messages from the study, and 29 (1.37%) did not respond to that question. A final model was run with the same unadjusted functional form as Model 1, but limited to participants who answered "Yes" or were imputed to have answered "Yes." The results are displayed in Table 3.

It is unknown why 459 of the 2115 participants (21.70%) reported not receiving messages from the study. Nonexclusive possibilities include phone numbers being miswritten on the survey, phone numbers being misentered into the SMS delivery system, participant phone numbers changing after enrollment, participants giving numbers beside their text-enabled cellular numbers as requested, and recall error.

This subset that reported receiving this study's texts displayed a stronger unadjusted intervention effect than that for all study participants. In this subset (Figure 1), the all texts group was associated with a highly significant reduction in the odds of undergoing CS (OR 0.66, *P*=.008).

Table 3. Logistic regression analysis of cesarean birth according to treatment assignment (women who reported text receipt).

Treatment assignment (unadjusted logistic regression)	Odds ratio	95% CI	*P* value
Basic only	Base case	N/A[a]	N/A
Care seeking	0.76	0.55-1.04	.09
Good home practices	0.80	0.58-1.09	.12
All texts	0.66	0.49-0.90	.008

[a]N/A: not applicable.

Discussion

External Validity

After adjusting for potentially confounding covariates, our main finding shows that the group with all texts sent together had an OR of 0.67 (*P*=.01). The rate of CS in the basic (comparison) group of our study (28.7% between 2013 and 2016) is identical to that in a study in rural Shaanxi, China in 2018 [36] and is well representative of (slightly lower than) the typical rural county rate of just over 30% in China in 2014, which was reported by Li et al [6]. Nonetheless, it is roughly double the WHO's recommended rate of 10% to 15%, and merits improvement. One possible reason for this disparity is that our sample came from an MCHC in a rural area. It has been estimated that women in large cities in China have 2.4 times the odds of CS as compared with women in smaller cities [37]. Further, Liu et al found that CS rates increased with hospital complexity, with tertiary care facilities having the highest rates [4]. Thus, our findings are likely to be most closely replicated in rural areas of China, but may be well indicative of such counties.

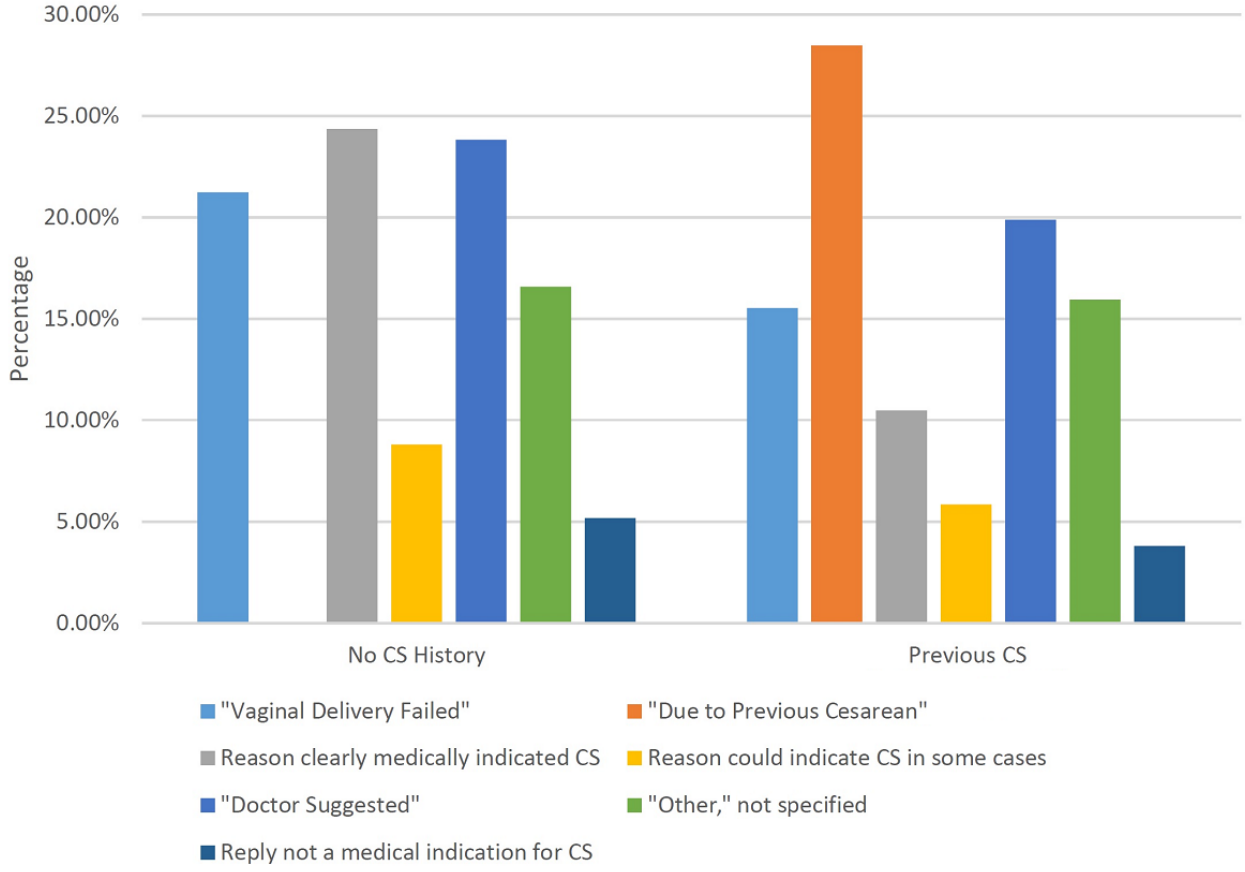
Potential Pathways

Our informational texts had several hypothesized pathways by which they might impact mode of delivery, including (1) altering women's underlying preferences, (2) empowering women to decline doctor-suggested CS by increasing their knowledge and self-efficacy, and/or (3) reducing the number of legitimate indications for CS. Our follow-up survey asked women who delivered via CS why they did so, and investigating their responses shed some light on how our intervention may have been effective. Due to the very strong association between a reported previous CS and mode of delivery, we first present responses broken down by CS history in Figure 2.

There was an unadjusted increase in the probability of delivering via CS of 54 percentage points, from about 27% probability among women without a previous CS to about 81% probability given a previous CS. A previous CS is often considered its own indication for CS and accounted for the largest difference between the two groups, but women with a history of CS had more CS procedures for every single reason category. The percentages in Figure 2 were derived from compiling the imputed "complete" data sets. The raw unimputed data had a smaller sample but indicated an even more stark difference in group rates of CS. Because of this strong association, we hypothesize that CS history could act as an effect modifier for our SMS intervention, but we were not powered to test this hypothesis directly. Therefore, we present in Multimedia Appendix 5 the stated reasons for CS by intervention group only for those women with no stated history of CS, accounting for about 87% of our sample. For convenience, Multimedia Appendix 6 presents the same data in terms of each arm's differences from the "basic" texts group.

CS based only on maternal request was not a common occurrence in our sample (Figure 2). Stated reasons for CS that were not a medical indication accounted for less than 5% of CS cases (26/548 cases) in our sample. This left our intervention very little room to affect CS by reducing cesarean delivery on maternal request, though the "other/not specified" category may or may not be comprised largely of maternal requests. Multimedia Appendix 6 shows that reducing clear medical indications for CS seems to be the strongest potential pathway for our texts to have had an effect, with this category showing the largest reduction in all treatment groups. The provider-induced "doctor-suggested" category had mixed evidence, being reduced in the "all texts" group but ostensibly increased in the individual "care-seeking" or "home practices" group.

Figure 2. Reasons for current cesarean section (CS) delivery by CS history.



Limitations

Our study has several limitations. First, women self-reported their delivery method to health workers. If our intervention induced a belief that vaginal delivery was a more socially desirable answer than CS, CS could be underreported in intervention arms. Second, our study population was concentrated in a single county. It is unknown how results would differ in other settings. Finally, high loss to follow-up lowered the planned statistical power of our study. Attrition does not seem to be associated with treatment assignment, and baseline covariates were balanced within both the full sample and the subset that completed the study, which suggests that loss to follow-up did not bias our results. However, we cannot measure whether it is associated with the mode of delivery, and as such, we cannot confidently rule out the possibility that high attrition altered our findings.

Implications

As a first of its kind, this evaluation breaks ground in the fields of SMS text messaging for maternal health in China and SMS text messaging for influencing the mode of delivery. In 2015, China had 16.55 million new births [38]. With the recent relaxations in China's one-child policy, this number could grow considerably in the next few years. Given the risk that unnecessary CS poses and the current excessive amount in China, wider distribution of the Newborn Health Project's messages on delivery mode seems to be a strategy worth trying. The acceptability and effectiveness of WeChat as a mode of delivering such messages also warrant scientific exploration.

Conclusions

A quasirandomized controlled trial distributing informational text messages to pregnant women in Gaoling, China found evidence that the full set of text messages may have reduced the number of cesarean deliveries in that group by 5.2 percentage points compared with the comparison group. Focusing on the subset of women who reported actually receiving program texts and adjusting for baseline covariates greatly strengthened this measured relationship. Given the numerous calls for strategies to reduce the rate of medically unnecessary CS in China, exploration of the wider distribution of these text messages seems warranted.

Acknowledgments

We would like to thank our implementation team members from Xi'an Jiaotong University. We are also grateful for all the support from leaders and health workers in the local collaborating organizations in Gaoling and Lantian. The authors would like to thank Drs David Canning, Till Bärnighausen, and Margaret McConnell for feedback on this manuscript. This study was made possible

with generous funding from the UBS Optimus Foundation and B&P Foundation. The funders had no role in the study except providing financial support.

Authors' Contributions

Yan S as the Project Director for the team led designing the intervention and study, led literature review, and detailed the method of program evaluation. JH participated in literature review, participated in development of survey questionnaires regarding psychological and behavioral measurements, led the data analysis, and led drafting of the manuscript. CY helped design the study, participated in literature review, and led the development of the text messages and survey questionnaires. Ya S, DW, and Zhi Z assisted in implementation, data entry, and data cleaning. Zho Z helped design the study and led the implementation of the project. All authors have made significant intellectual or practical contributions. All authors critically revised the manuscript and approved the final version of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
SMS text messages by general topic, treatment group, and timing.
[PDF File (Adobe PDF File), 87 KB - mhealth_v8i11e19953_app1.pdf]

Multimedia Appendix 2
Messages regarding delivery advice by treatment arm.
[PDF File (Adobe PDF File), 120 KB - mhealth_v8i11e19953_app2.pdf]

Multimedia Appendix 3
Balance check and all baseline variables: all enrollees.
[PDF File (Adobe PDF File), 405 KB - mhealth_v8i11e19953_app3.pdf]

Multimedia Appendix 4
Balance check and all baseline variables: women with follow-up surveys.
[PDF File (Adobe PDF File), 404 KB - mhealth_v8i11e19953_app4.pdf]

Multimedia Appendix 5
Reasons for current cesarean section delivery by SMS intervention assignment (N=2104).
[PDF File (Adobe PDF File), 11 KB - mhealth_v8i11e19953_app5.pdf]

Multimedia Appendix 6
Differences in reasons for current cesarean section delivery by SMS intervention assignment.
[PDF File (Adobe PDF File), 11 KB - mhealth_v8i11e19953_app6.pdf]

Multimedia Appendix 7
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2716 KB - mhealth_v8i11e19953_app7.pdf]

References

1. Vogel J, Betrán AP, Vindevoghel N, Souza J, Torloni M, Zhang J, WHO Multi-Country Survey on MaternalNewborn Health Research Network. Use of the Robson classification to assess caesarean section trends in 21 countries: a secondary analysis of two WHO multicountry surveys. Lancet Glob Health 2015 May;3(5):e260-e270 [FREE Full text] [doi: 10.1016/S2214-109X(15)70094-X] [Medline: 25866355]
2. WHO Statement on Caesarean Section Rates Internet. World Health Organization. 2015. URL: http://apps.who.int/iris/bitstream/10665/161442/1/WHO_RHR_15.02_eng.pdf [accessed 2020-10-25]
3. Lumbiganon P, Laopaiboon M, Gülmezoglu A, Souza J, Taneepanichskul S, Ruyan P, et al. Method of delivery and pregnancy outcomes in Asia: the WHO global survey on maternal and perinatal health 2007–08. The Lancet 2010 Feb;375(9713):490-499 [FREE Full text] [doi: 10.1016/s0140-6736(09)61870-5]
4. Liu Y, Li G, Chen Y, Wang X, Ruan Y, Zou L, et al. A descriptive analysis of the indications for caesarean section in mainland China. BMC Pregnancy Childbirth 2014 Dec 12;14:410 [FREE Full text] [doi: 10.1186/s12884-014-0410-2] [Medline: 25495550]

5. Betrán AP, Ye J, Moller A, Zhang J, Gülmezoglu AM, Torloni M. The Increasing Trend in Caesarean Section Rates: Global, Regional and National Estimates: 1990-2014. PLoS One 2016;11(2):e0148343-e0142014 [FREE Full text] [doi: 10.1371/journal.pone.0148343] [Medline: 26849801]
6. Li H, Luo S, Trasande L, Hellerstein S, Kang C, Li J, et al. Geographic Variations and Temporal Trends in Cesarean Delivery Rates in China, 2008-2014. JAMA 2017 Jan 03;317(1):69-76. [doi: 10.1001/jama.2016.18663] [Medline: 28030701]
7. Villar J, Carroli G, Zavaleta N, Donner A, Wojdyla D, Faundes A, World Health Organization 2005 Global Survey on MaternalPerinatal Health Research Group. Maternal and neonatal individual risks and benefits associated with caesarean delivery: multicentre prospective study. BMJ 2007 Nov 17;335(7628):1025 [FREE Full text] [doi: 10.1136/bmj.39363.706956.55] [Medline: 17977819]
8. American College of ObstetriciansGynecologists (College), Society for Maternal-Fetal Medicine, Caughey AB, Cahill AG, Guise JM, Rouse DJ. Safe prevention of the primary cesarean delivery. Am J Obstet Gynecol 2014 Mar;210(3):179-193. [doi: 10.1016/j.ajog.2014.01.026] [Medline: 24565430]
9. Liu X, Landon MB, Cheng W, Chen Y. Cesarean delivery on maternal request in China: what are the risks and benefits? Am J Obstet Gynecol 2015 Jun;212(6):817.e1-817.e9 [FREE Full text] [doi: 10.1016/j.ajog.2015.01.043] [Medline: 25640048]
10. Visco AG, Viswanathan M, Lohr KN, Wechter ME, Gartlehner G, Wu JM, et al. Cesarean Delivery on Maternal Request. 2006;108(6):1517-1529. [doi: 10.1097/01.aog.0000241092.79282.87]
11. Souza J, Gülmezoglu AM, Lumbiganon P, Laopaiboon M, Carroli G, Fawole B, WHO Global Survey on MaternalPerinatal Health Research Group. Caesarean section without medical indications is associated with an increased risk of adverse short-term maternal outcomes: the 2004-2008 WHO Global Survey on Maternal and Perinatal Health. BMC Med 2010 Nov 10;8:71-7015 [FREE Full text] [doi: 10.1186/1741-7015-8-71] [Medline: 21067593]
12. Thompson J, Roberts C, Currie M, Ellwood D. Prevalence and persistence of health problems after childbirth: associations with parity and method of birth. Birth 2002 Jun;29(2):83-94 [FREE Full text] [doi: 10.1046/j.1523-536x.2002.00167.x] [Medline: 12051189]
13. Kainu J, Sarvela J, Tiippana E, Halmesmäki E, Korttila K. Persistent pain after caesarean section and vaginal birth: a cohort study. Int J Obstet Anesth 2010 Jan;19(1):4-9 [FREE Full text] [doi: 10.1016/j.ijoa.2009.03.013] [Medline: 19733050]
14. Gilliam M. Cesarean delivery on request: reproductive consequences. Semin Perinatol 2006 Oct;30(5):257-260. [doi: 10.1053/j.semperi.2006.07.005] [Medline: 17011396]
15. Feng XL, Wang Y, An L, Ronsmans C. Cesarean section in the People's Republic of China: current perspectives. Int J Womens Health 2014;6:59-74 [FREE Full text] [doi: 10.2147/IJWH.S41410] [Medline: 24470775]
16. Zhang J, Liu Y, Meikle S, Zheng J, Sun W, Li Z. Cesarean delivery on maternal request in southeast China. Obstet Gynecol 2008 May;111(5):1077-1082. [doi: 10.1097/AOG.0b013e31816e349e] [Medline: 18448738]
17. Huang C, Wu H, Huang S, Chien L, Guo J. Transtheoretical model-based passive smoking prevention programme among pregnant women and mothers of young children. Eur J Public Health 2013 Oct 15;23(5):777-782 [FREE Full text] [doi: 10.1093/eurpub/cks177] [Medline: 23325124]
18. Mi J, Liu F. Rate of caesarean section is alarming in China. The Lancet 2014 Apr;383(9927):1463-1464 [FREE Full text] [doi: 10.1016/s0140-6736(14)60716-9]
19. Ji H, Jiang H, Yang L, Qian X, Tang S. Factors contributing to the rapid rise of caesarean section: a prospective study of primiparous Chinese women in Shanghai. BMJ Open 2015 Nov 13;5(11):e008994-e008994. [doi: 10.1136/bmjopen-2015-008994] [Medline: 26567254]
20. Noordam A, Kuepper B, Stekelenburg J, Milen A. Improvement of maternal health services through the use of mobile phones. Trop Med Int Health 2011 May;16(5):622-626 [FREE Full text] [doi: 10.1111/j.1365-3156.2011.02747.x] [Medline: 21342374]
21. Tamrat T, Kachnowski S. Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world. Matern Child Health J 2012 Jul;16(5):1092-1101. [doi: 10.1007/s10995-011-0836-3] [Medline: 21688111]
22. Evans WD, Wallace JL, Snider J. Pilot evaluation of the text4baby mobile health program. BMC Public Health 2012 Nov 26;12(1):1031 [FREE Full text] [doi: 10.1186/1471-2458-12-1031] [Medline: 23181985]
23. Jareethum R, Titapant V, Chantra T, Sommai V, Chuenwattana P, Jirawan C. Satisfaction of healthy pregnant women receiving short message service via mobile phone for prenatal support: A randomized controlled trial. J Med Assoc Thai 2008 Apr;91(4):458-463. [Medline: 18556852]
24. Lund S, Hemed M, Nielsen B, Said A, Said K, Makungu M, et al. Mobile phones as a health communication tool to improve skilled attendance at delivery in Zanzibar: a cluster-randomised controlled trial. BJOG 2012 Sep;119(10):1256-1264 [FREE Full text] [doi: 10.1111/j.1471-0528.2012.03413.x] [Medline: 22805598]
25. Naughton F, Jamison J, Sutton S. Attitudes towards SMS text message smoking cessation support: a qualitative study of pregnant smokers. Health Educ Res 2013 Oct 02;28(5):911-922 [FREE Full text] [doi: 10.1093/her/cyt057] [Medline: 23640985]

26. Tang S, Ghose B, Hoque MR, Hao G, Yaya S. Women Using Mobile Phones for Health Communication Are More Likely to Use Prenatal and Postnatal Services in Bangladesh: Cross-Sectional Study. JMIR Mhealth Uhealth 2019 Feb 28;7(2):e10645 [FREE Full text] [doi: 10.2196/10645] [Medline: 30816850]
27. Hussain T, Smith P, Yee LM. Mobile Phone-Based Behavioral Interventions in Pregnancy to Promote Maternal and Fetal Health in High-Income Countries: Systematic Review. JMIR Mhealth Uhealth 2020 May 28;8(5):e15111 [FREE Full text] [doi: 10.2196/15111] [Medline: 32463373]
28. Zhou Z, Su Y, Heitner J, Si Y, Wang D, Zhou Z, et al. The Effects on Inappropriate Weight for Gestational Age of an SMS Based Educational Intervention for Pregnant Women in Xi'an China: A Quasi-Randomized Controlled Trial. Int J Environ Res Public Health 2020 Feb 25;17(5):1482 [FREE Full text] [doi: 10.3390/ijerph17051482] [Medline: 32106558]
29. Su Y, Yuan C, Zhou Z, Heitner J, Campbell B. Impact of an SMS advice programme on maternal and newborn health in rural China: study protocol for a quasi-randomised controlled trial. BMJ Open 2016 Aug 10;6(8):e011016. [doi: 10.1136/bmjopen-2015-011016] [Medline: 27515750]
30. Pannel VM. Miscarriages: Diagnosis, Management and Coping Strategies. Hauppauge, NY: Nova Science Publishers; 2014. URL: https://novapublishers.com/shop/miscarriages-diagnosis-management-and-coping-strategies/
31. Wang C. Induced abortion patterns and determinants among married women in China: 1979 to 2010. Reproductive Health Matters 2014 Jun 04;22(43):159-168. [doi: 10.1016/s0968-8080(14)43753-4]
32. King G, Honaker J, Joseph A, Scheve K. Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation. Am Polit Sci Rev 2002 Mar 28;95(1):49-69 [FREE Full text] [doi: 10.1017/s0003055401000235]
33. de Jersey SJ. Are we missing opportunities? Understanding health behaviours associated with overweight in pregnancy. QUT ePrints. 2013. URL: https://eprints.qut.edu.au/61090/ [accessed 2020-10-25]
34. Lippke S, Ziegelmann JP. Theory-based health behavior change: Developing, testing, and applying theories for evidence-based interventions. Applied Psychology: An International Review 2008;57(4):698-716. [doi: 10.1111/j.1464-0597.2008.00339.x]
35. Noar SM, Zimmerman RS. Health Behavior Theory and cumulative knowledge regarding health behaviors: are we moving in the right direction? Health Educ Res 2005 Jun;20(3):275-290 [FREE Full text] [doi: 10.1093/her/cyg113] [Medline: 15632099]
36. Zhang R, Li S, Li C, Zhao D, Guo L, Qu P, et al. Socioeconomic inequalities and determinants of maternal health services in Shaanxi Province, Western China. PLoS One 2018;13(9):e0202129 [FREE Full text] [doi: 10.1371/journal.pone.0202129] [Medline: 30183720]
37. Tang S, Li X, Wu Z. Rising cesarean delivery rate in primiparous women in urban China: evidence from three nationwide household health surveys. Am J Obstet Gynecol 2006 Dec;195(6):1527-1532 [FREE Full text] [doi: 10.1016/j.ajog.2006.03.044] [Medline: 16677593]
38. China's Economy Realized a Moderate but Stable and Sound Growth in 2015. National Bureau of Statistics of China. 2016. URL: http://www.stats.gov.cn/english/PressRelease/201601/t20160119_1306072.html [accessed 2020-10-25]

Abbreviations

CS: cesarean section
MCHC: maternal and child health center
mHealth: mobile health
OR: odds ratio
WHO: World Health Organization

Edited by G Eysenbach; submitted 07.05.20; peer-reviewed by P Fontelo, E Luo; comments to author 20.08.20; revised version received 12.09.20; accepted 22.09.20; published 03.11.20.

Please cite as:
Su Y, Heitner J, Yuan C, Si Y, Wang D, Zhou Z, Zhou Z
Effect of a Text Messaging–Based Educational Intervention on Cesarean Section Rates Among Pregnant Women in China: Quasirandomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(11):e19953
URL: https://mhealth.jmir.org/2020/11/e19953
doi:10.2196/19953
PMID:33141099

Original Paper

Network Support Using Social Networking Services to Increase Exercise Adherence Among Korean-Chinese Middle-Aged Migrant Women: Mixed Methods Study

Hyeyeon Lee[1], MPH; Hyeonkyeong Lee[2], PhD; Youlim Kim[1], MPH; Sookyung Kim[1], MPH; Young-Me Lee[3], PhD

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

[2]Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

[3]School of Nursing, College of Science and Health, Depaul University, Chicago, IL, United States

Corresponding Author:
Hyeonkyeong Lee, PhD
Mo-Im Kim Nursing Research Institute
College of Nursing
Yonsei University
50-1 Yonse-ro, Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2228 3373
Fax: 82 2392 5440
Email: hlee39@yuhs.ac

Abstract

Background: Social networking services (SNSs) are recognized to be a promising approach to easily deliver health interventions and to enhance social support for exercise adherence. However, the patterns and aspects of social support through SNSs have not been reported and their influence on other social-cognitive factors remains inconclusive.

Objective: Our objective is to explore how social support delivered through SNSs impacts interactions among Korean-Chinese (KC) middle-aged women and to identify how this approach influences social-cognitive factors for exercise (eg, sense of community, self-efficacy for exercise, and social support for exercise).

Methods: A mixed methods design was used. Text analysis of SNS messages and text mining using the Korean Natural Language Application (KoALA) were conducted. Social-cognitive factors (eg, sense of community, self-efficacy for exercise, and social support for exercise) were assessed at baseline and after 12 weeks using a structured questionnaire. A comparison of social-cognitive factors at baseline and at 12 weeks was conducted to identify any potential significant changes, using the Wilcoxon signed-rank test.

Results: A total of 259 SNS messages were collected from 24 KC women, distributed among four chat groups, who participated in a 12-week walking intervention program between August and October 2018. The individual average frequency of chatting via the SNS was 10.79 (range 0-34) and the most frequent type of social support through the SNS was network support (172/259, 66.4%). The most common words extracted from the SNS were *Health*, *Exercise*, *Participation*, and *We*. Overall, the perceived levels of sense of community (P<.001) and social support for exercise (P=.002) were significantly increased at 12 weeks compared with baseline. Group 1 (P=.03) and Group 4 (P=.03), whose members demonstrated the highest frequency of network support, experienced a significant increase only in the level of sense of community.

Conclusions: By integrating these data and conducting a mixed methods analysis, we observed that among the types of social support, network support was a key point for the promotion of social-cognitive factors in increasing exercise adherence. Therefore, network support through SNS-based interventions should be considered as a useful strategy to help vulnerable migrant populations make changes to exercise behaviors.

(JMIR Mhealth Uhealth 2020;8(11):e19159) doi:10.2196/19159

KEYWORDS

SNS; social support; network support; exercise adherence; social-cognitive factors; text mining

Introduction

Social Networking Services as Useful Tools for Health Promotion

Interpersonal relationships and social interactions are shifting from face-to-face to online environments due to the development of the internet and smart devices, as well as an increase in economic development [1]. Social networking services (SNSs) have enabled individuals to interact and connect with others more easily, regardless of time and place, including those with diverse racial and ethnic backgrounds. While there are negative influences of SNSs on well-being due to social comparison and isolation [2], individuals easily establish new interpersonal networks and experience psychological well-being (eg, happiness and satisfaction with life) by using SNSs, such as Facebook or Twitter [3,4]. SNSs are also recognized as valuable and excellent platforms for delivering health interventions to hard-to-reach minority populations [4].

Exercise and Social-Cognitive Factors in Migrant Populations

The population of South Korea is multicultural, with more than 2.1 million migrants in 2017, including roughly 680,000 (32.9%) Korean-Chinese (KC) [5]. The proportion of this KC population that is made up of middle-aged women (ages 40-65 years) is 33.17% (n=225,488), and these women often work in the restaurant service industry or as houseworkers [5,6]. Due to changes in their health, middle-aged women are vulnerable to physical and mental health problems (eg, cardiovascular disease, type 2 diabetes, obesity, hypertension, and depression) [7,8]. Although physical activities are important for managing the health of middle-aged women, migrant groups often have lower participation rates in programs that promote a healthy lifestyle compared to other groups, due to cultural differences, insufficient knowledge of health services, and limited access [9,10]; access may be considerably limited due to the working environment (eg, less flexible and less autonomous) compared with native Koreans [10]. These women face many other barriers, including physical-related barriers (eg, lack of time and space to exercise), knowledge-related barriers (eg, lack of knowledge about exercise), psychological barriers (eg, low self-efficacy), and lack of social support [11]. As a result, it is especially difficult to implement physical activity–based interventions for middle-aged migrant women.

As a strategy to encourage exercise adherence in migrant populations, social-cognitive factors such as social support have been commonly adopted in interventional designs [12]. Recently, SNS-based interventions emphasizing social support as well as social-cognitive factors have been found to increase exercise adherence. For instance, in one study of adults with type 2 diabetes, participants in a private Facebook group that received feedback, coaching, and social support in real time increased the number of steps taken in a day compared to those in the nonintervention group [13]. However, earlier studies only identified and categorized the perceived degree of social support through SNSs without an analysis of the patterns or aspects of social support during the intervention period [14,15]. In addition, although existing evidence demonstrates that social support has a positive correlation with other social-cognitive factors (eg, sense of community and self-efficacy for exercise) [16-18], there are still a limited number of studies, and results remain inconclusive.

Definition and Types of Social Support

Overview

Social support is defined as interpersonal interactions that involve emotional attention, instrumental assistance, information about the environment, and a positive self-assessment [19,20]. Types of social support through SNSs in this study are based on previous studies and include the following: (1) network support, (2) emotional support, (3) information support, and (4) esteem support [19-21]. Exceptionally tangible or instrumental support was excluded from the analysis because it is based on private topics and offline conversations [22]. Additionally, gratitude is related to social support and substantially and positively affects emotional support; this study included the expression of gratitude as a subdomain of emotional support [23].

Network Support

Network support is defined as the presence of companions or potential aides (eg, “Find assistants to help you”) who share or engage in social activities and resources [2,21,24]. Network support is identified from the desire to be emotionally satisfied and to maintain a relationship with a group to which one belongs [25]. In the online health community, a network could include conversations about members’ interests [21] and life-related topics that are not related to health issues [22,26]. Subcategories of network support are access, presence, and companionship [20,27]. Access refers to a person in a similar situation asking questions on behalf of you or sharing mailing address and/or contact information. Presence involves a supportive expression indicating that you are willing to help others. Companionship includes messages that reassure members that they share a similar environment and experience, including experiences not related to health issues [28] (eg, nonverbal cues, such as emoticons, narratives, poetry, humor, jokes, and vacation planning).

Emotional Support

Emotional support includes expression, encouragement, sympathy, affection, empathy, care, and concern [14,22,26] (eg, “You can do it!” “I believe,” and “Keep it up”). Subcategories of emotional support are relationship, physical affection, confidentiality, sympathy, understanding or empathy, encouragement, prayer, and gratitude. Relationships are indicated with important messages that express friendliness or closeness within the online community. “Good morning” and “Nice to meet you” are often used as the first phrases of conversation conveying emotional support. Physical affection in an SNS is delivered as text-based messages (eg, “HUG”) and provides emotional support, such as physical stability. Confidentiality is related to protection of secrets, whereas sympathy is an expression of sorrow for another member’s stress and suffering. Understanding is an expression of empathy to another member who is experiencing difficult and/or complex situations. Encouragement includes messages that convey hope

and confidence to other members. Prayer offers a blessing for other members, and gratitude is an expression of appreciation.

Information Support

Information support involves the delivery of information or the provision of guidelines or suggestions (eg, information about physical activity, recommendations, and how to deal with symptoms) [14,22,26]. Subcategories of information support include advice, referral, situational appraisal, and teaching. Referral involves sharing specific informational resources (eg, books, websites, and institutions) with other members. Situational appraisal might come in the form of a message to help re-evaluate another member's situation. Teaching involves the dissemination of factual information.

Esteem Support

Esteem support involves sharing a complimentary expression, such as "Congratulations" or "Great job" [14,26]. Subcategories of esteem support are compliment, validation, and relieving of blame. A compliment is a celebration of accomplishment despite a negative situation or a positive statement to other members. Validation involves agreeing with the feelings and stories of other members or to support their view of the situation. Relieving of blame comes in the form of an expression of support for another member who expresses guilt for a particular situation.

Purpose of This Study

The purpose of this study was to (1) explore how social support through SNSs appeared in interactions among KC middle-aged women who are vulnerable to health problems and (2) identify how this social support influenced social-cognitive factors for exercise (eg, sense of community, self-efficacy for exercise, and social support for exercise).

Methods

Design

The mixed methods design was used to combine qualitative and quantitative data to better understand patterns of social support and its effect on other social-cognitive factors potentially impacting exercise adherence.

Participants and Data Collection

A community-based project was conducted to improve physical and mental health of KC migrant women through a 6-month, culturally adaptive, mobile app–based walking intervention between May 2017 and February 2019 [29]. For recruiting participants, we visited a public health center, the Korea Support Center for Foreign Workers, and a KC church; distributed leaflets; and promoted the health intervention program. Participants in the study were KC middle-aged women (40-65 years of age) who had no health restrictions against increasing physical activity. Based on physical activity assessed using the Physical Activity Readiness Questionnaire, those who walked more than three times a week and more than 30 minutes at a time were excluded. Only those who were able to use the smartphone app and participate in the program for 6 months were included in the study. There were 28 participants in the experimental group, all of whom completed the baseline assessment after completing the informed consent form. In order to promote social support and networking for physical activity, these 28 KC middle-aged women were divided into four groups. A group leader was appointed by vote within each group and instructed to encourage and support other group members throughout the study to increase exercise adherence.

Data collection was conducted over an intervention period of 12 weeks between August and October 2018. Out of the 28 participants, 1 (4%) dropped out at the fifth week and 2 (7%) did not complete the 12-week assessment and were, thus, excluded from analysis. During the 12-week intervention period, a total of 353 messages were collected via KakaoTalk, the most common chatting app in Korea. However, a woman whose comments made up a disproportionate number of total SNS conversations (94/353, 26.6%) was excluded because her contributions may have distorted the study results. Ultimately, all 259 text messages written in Korean from the SNS groups of 24 KC middle-aged women were used in the final analysis of this study.

Qualitative Data Analysis: Text Analysis and Text Mining

Text Analysis

Text analysis was conducted on 259 messages sent among the 24 participants via SNS over the 12 weeks by types of social support. Two researchers (Hyeyeon L and YK) independently coded all 259 messages according to social support codes. Inconsistent codes were reviewed and discussed with the principal investigator (Hyeonkyeong L) to reach consensus [14].

Text Mining

Text mining was conducted using the Korean Natural Language Application (KoALA) (Mondata Co) to identify themes and interests of KC middle-aged women interacting using an SNS. The KoALA engine is an application for social media mining and supports the entire text-mining process, including social media data collection, preprocessing, and analysis [30]. A word cloud visually highlights keywords based on how frequently they are used. Co-occurrence analysis is a matrix form that calculates the weight based on the frequency of keywords appearing in the text at the same time. The simultaneous appearance of certain keywords in the text means that they are related to each other, and the repetition of these keywords indicates a high correlation between the keywords.

Four chat room conversations undertaken over 12 weeks using an SNS were integrated. After extracting all text and dividing it into sentence units, the process of recognizing entity names was conducted using the entity name dictionary. The next step was to remove stop words. Data analysis was performed using the word cloud approach and co-occurrence matrix. After preprocessing was complete, 20 nouns were extracted from the chat group in order of appearance. Those were used to analyze the word *exercise* and the words that appeared as co-occurring terms.

Quantitative Data Analysis

Overview

Social-cognitive factors (eg, sense of community, self-efficacy for exercise, and social support for exercise) were assessed at baseline and at 12 weeks. General characteristics, mean, standard deviation, and frequency of data were analyzed by descriptive statistics using SPSS Statistics for Windows, version 25.0 (IBM Corp). Additionally, the Wilcoxon signed-rank test was conducted to confirm the significance of sense of community, self-efficacy for exercise, and social support for exercise by individuals and groups.

Sense of Community

Sense of community was measured using the Korean version of the Brief Sense of Community Scale, a tool based on the scale developed by Peterson et al [31] and later modified by Oh [32]. The scale involves eight items rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). A higher score represents a greater sense of community. Among the items, *peer/friend* was changed to *neighbor*, and *peer group* was changed to *neighborhood*. In the study, the Cronbach alpha was .73.

Self-Efficacy for Exercise

Self-efficacy for exercise was measured using the Korean version of the Barrier Self-Efficacy Scale developed by McAuley [33] and modified by Choi [34]. The scale includes 13 items, and ratings range from 0 (completely certain that I could not) to 10 (completely certain that I could). The higher the average score, the higher the self-efficacy for exercise. In the study, the Cronbach alpha was .89.

Social Support for Exercise

Social support for exercise was assessed using a questionnaire developed by Sallis et al [35] and translated and modified by Choi [36]. In this study, six items were selected, except the one added by Choi. The scale includes six items that assess perceived social support for exercise from family and friends during the past 3 months. The scale is rated using a 4-point Likert scale, ranging from 1 (strongly disagree) to 4 (strongly agree). The higher the total score, the higher the social support from family and friends. In the study, the Cronbach alpha was .88.

Results

Participant Characteristics

A total of 24 KC middle-aged women participated in the SNS groups. The average age of participants was 47.2 (SD 6.5) years; average time spent working daily was 10 (SD 4) hours; average duration of stay in Korea was 144.2 (SD 79.1) months (about 12 years); average number of months working at current job was 39.6 (SD 50.3); and average monthly income was about US $1504.50 (SD 469.75). Most participants had an educational level above high school graduate (17/24, 71%), and the service industry (ie, cook, waiter, housekeeper, hairdresser, and caregiver) was the most common job type (10/24, 42%). Moreover, 259 text messages written in Korean from the SNS groups were collected over an intervention period of 12 weeks between August and October 2018. The individual average frequency of chats via SNS was 10.79 (range 0-34).

Qualitative Findings

Text Analysis: Social Support in the SNS

Social support in this study was divided into 18 subcategories (see Table 1). The most frequent type of social support through the SNS identified by this group was network support (172/259, 66.4%), followed by emotional support (40/259, 15.4%), information support (28/259, 10.8%), and esteem support (19/259, 7.3%).

Table 1. Conceptual definitions of social support and frequency of messages of each type.

Categories and subcategories of social support	Frequency of messages (N=259), n (%)
Network support	
Total	172 (66.4)
1. Access	0 (0)
2. Presence	0 (0)
3. Companionship	172 (66.4)
Emotional support	
Total	40 (15.4)
4. Relationship	11 (4.2)
5. Physical or visual affection	2 (0.8)
6. Confidentiality	0 (0)
7. Sympathy	1 (0.4)
8. Understanding or empathy	0 (0)
9. Encouragement	14 (5.4)
10. Prayer	0 (0)
11. Gratitude	12 (4.6)
Information support	
Total	28 (10.8)
12. Advice	10 (3.8)
13. Referral	0 (0)
14. Situational appraisal	1 (0.4)
15. Teaching	17 (6.6)
Esteem support	
Total	19 (7.3)
16. Compliment	13 (5.0)
17. Validation	6 (2.3)
18. Relieving of blame	0 (0)

In total, 66.4% (172/259) of the support messages collected from the SNS were network support messages. Network support messages made up three of the 18 subcategories of social support; of these three subcategories, all messages involved companionship (172/172, 100%). Companionship includes a response, for example, to attending an offline cultural class, naming a group, talking about daily life, and discussion about exercise barriers; companionship may come in the form of short opinions, photos, and emoticons (eg,  and ). No messages corresponded to access or presence.

Emotional support messages within the SNS made up 15.4% (40/259) of all support messages and eight of the 18 subcategories of social support. In this study, all emotional support messages fell into one of five of these eight subcategories: (1) encouragement (14/259, 5.4%), (2) gratitude (12/259, 4.6%), (3) relationship (11/259, 4.2%), (4) physical affection (2/259, 0.8%), and (5) sympathy (1/259, 0.4%). No messages corresponding to confidentiality, understanding or empathy, or prayer were noted.

Information support messages within the SNS made up 10.8% (28/259) of all support messages and four of the 18 social support subcategories. In this study, all information support messages fell into one of three of these four subcategories: (1) teaching (17/259, 6.6%), (2) advice (10/259, 3.8%), and (3) situational appraisal (1/259, 0.4%). No referral messages were observed in this study.

Esteem support messages obtained from the SNS accounted for 7.3% (19/259) of all messages. These esteem support messages are categorized into two out of three of the 18 subcategories of social support: (1) compliment (13/259, 5.0%) and (2) validation (6/259, 2.3%). No messages from this study were categorized as relieving of blame. Example quotes by the type of social support are addressed in Table 2.

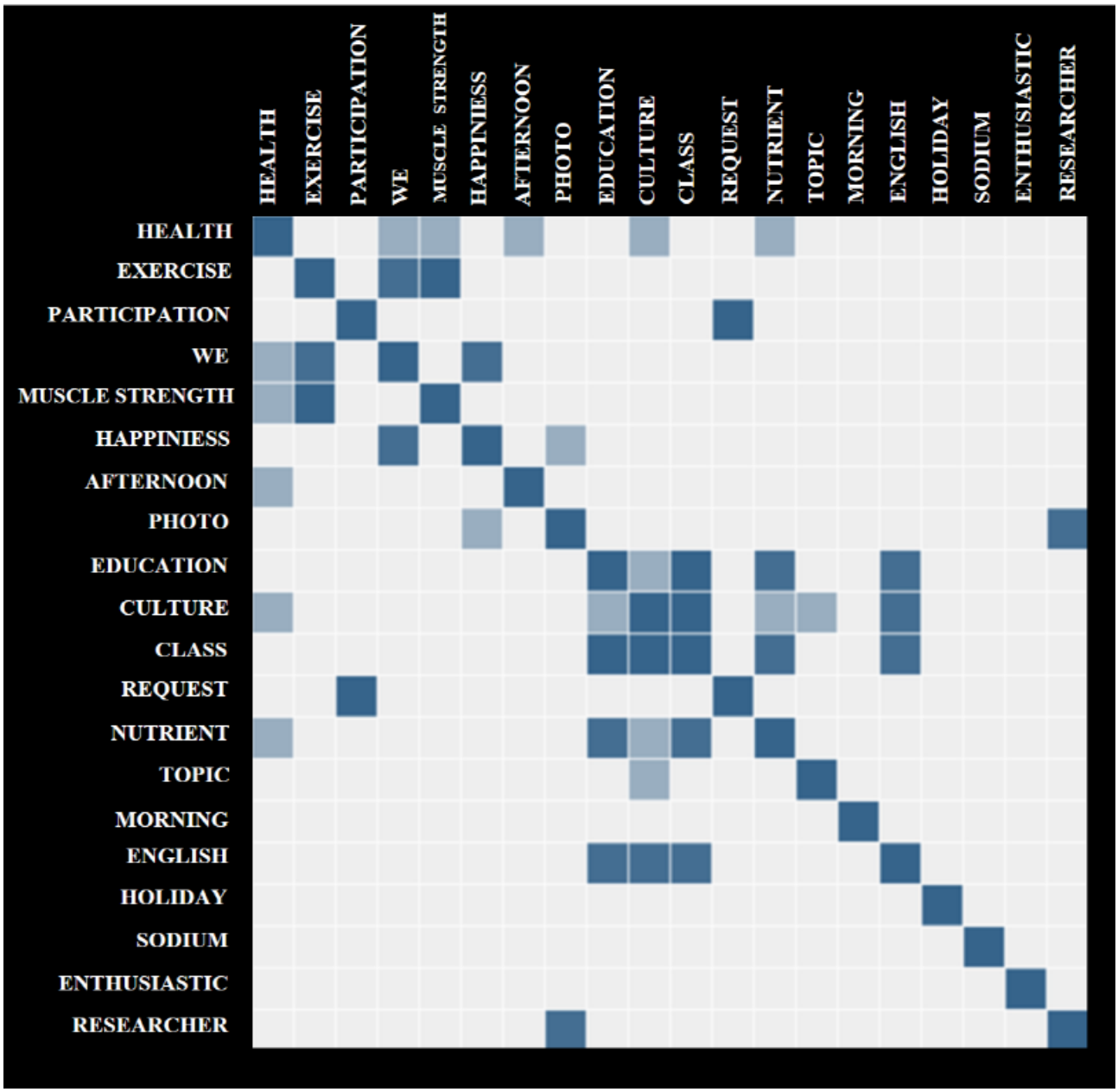
Table 2. Example quotes from messages of each type of social support.

Type of support	Quote
Network support: companionship	
Offline culture class	"It is difficult for me to attend." [ID 5, September 9, Group 1]
	"Ok, I will attend." [ID 9, August 15, Group 1]
	"Where will it be held?" [ID 2, September 4, Group 1]
Naming a group	"Beautiful Women Club?" [ID 45, August 17, Group 4]
	"How about Happy Club?" [ID 26, August 10, Group 3]
	"I'd like to recommend 'Happy Club' too; I think if we are not happy, our health is meaningless." [ID 28, August 14, Group 3]
Daily life	"I'm on vacation for 5 days." [ID 2, August 21, Group 1]
	"I'll have salad for breakfast." [ID 2, September 14, Group 1]
	"What sauce?" [ID 5, September 14, Group 1]
Exercise barriers	"I do well walking, but muscle exercise is difficult for me." [ID 20, August 10, Group 2]
	"It is hard to keep myself motivated for doing muscle exercise." [ID 18, August 10, Group 2]
	"I can't exercise because the weather is hot." [ID 25, August 10, Group 2]
Emotional support	
Relationship	"Nice to meet you." [ID 22, August 10, Group 2]
	"Have a nice day." [ID 2, September 17, Group 1]
	"Have a nice weekend." [ID 26, September 29, Group 4]
Physical affection	"Emoticon: kiss." [ID 2, August 29, Group 1]
Sympathy	"I'm sorry to hear that." [ID 26, September 4, Group 3]
Encouragement	"Let's exercise this week as well for health." [ID 43, August 23, Group 4]
Gratitude	"Thanks everyone." [ID 26, August 23, Group 3]
Information support	
Advice	"To eat less sodium, I put some vegetables in Korean noodles." [ID 19, September 18, Group 3]
Situational appraisal	"This offline cultural class is an essential education for a healthy life, I think." [ID 26, September 4, Group 3]
Teaching	"The phone will work again if you turn the phone off and on." [ID 18, August 13, Group 2]
	"It is good to leave three holes of band to wear the Fitbit band." [ID 22, August 10, Group 2]
Esteem support	
Compliment	"Congratulations." [ID 2, August 25, Group 1]
	"The yellow umbrella and your elegant fashion go well together; it's so pretty." [ID 26, August 28, Group 2]
Validation	"I agree." [ID 28, August 10, Group 3]
	"When you put on makeup after taking a makeup class, your face definitely looks better." [ID 28, August 20, Group 3]

Text-Mining Findings

A keyword-extraction algorithm was used to identify the top 20 keywords and confirm whether these keywords were related to the exercise-adherence program analyzed here. The most common words extracted from the SNS were *Health* (n=21), *Exercise* (n=20), *Participation* (n=18), and *We* (n=18). The keyword used the most was *Health*, which confirmed that participants mainly talked about health-related topics on the SNS. The second-most used keyword was *Exercise*. Participants talked about their compliance with muscular exercise and walking related to *Exercise*. The third-most used keywords were *Participation* and *We*, confirming that these 24 KC middle-aged women encouraged one another to participate in exercise, referring to themselves as "We" with the goal of exercise adherence. As a result of co-occurrence analysis, the words that appeared to be highly related to *Health* were *Exercise*, *We*, *Muscle strength*, and *Happiness* (see Figure 1). It was found that KC middle-aged women encouraged *Exercise* and formed a sense of belonging of *We* regarding health in the SNS. The words *We* and *Happiness* were highly related, and the words that appeared most often with *Culture* were *Education* and *Class*.

Figure 1. Co-occurrence matrix. The darker the color of the square, the more frequently the pair of words appears.



Quantitative Findings

Changes in the Levels of Social-Cognitive Factors at Baseline and at 12 Weeks

The Wilcoxon signed-rank test was conducted to analyze changes in sense of community, self-efficacy for exercise, and social support for exercise among KC middle-aged women from baseline to 12 weeks. Overall, the perceived levels of sense of community (z=–3.30, P<.001) and social support for exercise (z=–3.09, P=.002) were significantly increased at 12 weeks compared with baseline (see Table 3). The level of sense of community after the 12-week intervention increased significantly only in Group 1 (z=–2.20, P=.03) and Group 4 (z=–2.20, P=.03). The only increase in self-efficacy for exercise was in Group 1 (z=–2.03, P=.04), and the only increase in social support for exercise was in Group 4 (z=–2.02, P=.04).

Table 3. Changes in the levels of social-cognitive factors at baseline and 12 weeks.

Factors	Total (N=24)		Group 1 (n=6)		Group 2 (n=8)		Group 3 (n=4)		Group 4 (n=6)	
	Mean (SD)	*P* value[a]	Mean (SD)	*P* value[a]	Mean (SD)	*P* value[a]	Mean (SD)	*P* value[a]	Mean (SD)	*P* value[a]
Sense of community[b]										
Baseline	23.79 (7.35)	*<.001* [c]	24.17 (3.66)	*.03*	21.75 (8.71)	.80	18.75 (7.46)	.07	29.50 (5.39)	*.03*
12 weeks	29.25 (8.24)		31.17 (4.02)		22.00 (7.60)		28.00 (7.35)		37.83 (2.14)	
Self-efficacy for exercise[d]										
Baseline	6.92 (1.95)	.18	6.17 (2.32)	*.04*	6.75 (1.58)	.23	8.00 (2.00)	.71	7.17 (2.14)	.07
12 weeks	7.46 (2.30)		8.67 (1.21)		5.63 (2.67)		7.75 (0.50)		8.50 (2.07)	
Social support for exercise[e]										
Baseline	14.83 (5.22)	*.002*	17.00 (3.63)	.21	13.13 (3.83)	.12	15.00 (7.75)	.29	14.83 (6.74)	*.04*
12 weeks	18.63 (4.17)		19.50 (2.59)		15.00 (4.01)		20.50 (1.92)		21.33 (3.78)	

[a]The Wilcoxon signed-rank test was conducted.

[b]Measured using the Korean version of the Brief Sense of Community Scale; scores range from 1 (strongly disagree) to 5 (strongly agree).

[c]Italicized *P* values indicate significant results.

[d]Measured using the Korean version of the Barrier Self-Efficacy Scale; scores ranged from 0 (completely certain that I could not) to 10 (completely certain that I could).

[e]Assessed using a questionnaire developed by Sallis et al [35] and translated and modified by Choi [36]; scores ranged from 1 (strongly disagree) to 4 (strongly agree).

The Frequency of Each Type of Social Support Message Provided by Each Group

After analyzing the social support pattern and frequency by group, the frequency of network support was found to be higher than other types of social support in all groups. When comparing the social support frequency by group, Groups 1 and 4 demonstrated the highest network support frequency of over 70% (see Table 4). Due to the limited amount of data, the correlation and significance among types of social support and social-cognitive factors could not be studied.

Table 4. The frequency of each type of social support message provided by each group.

Group	Social support messages by type (N=259), n (%)				
	Network	Emotional	Information	Esteem	Total
1	56 (73.7)	10 (13.2)	3 (3.9)	7 (9.2)	76 (29.3)
2	33 (62.3)	5 (9.4)	14 (26.4)	1 (1.9)	53 (20.5)
3	35 (48.6)	18 (25.0)	11 (15.3)	8 (11.1)	72 (27.8)
4	48 (82.7)	7 (12.1)	0 (0)	3 (5.2)	58 (22.4)
Total	172 (66.4)	40 (15.4)	28 (10.8)	19 (7.4)	259 (100)

Discussion

Principal Findings

The main result of this study is that the most frequent type of social support through the SNS was network support and that levels of social-cognitive factors, including sense of community, self-efficacy for exercise, and social support for exercise, increased in groups with a high frequency of network support. Moreover, the simultaneously appearing keywords were *We*, *Exercise*, and *Happiness*. This study analyzed SNS messages shared among KC middle-aged women who participated in mobile app–based health promotion programs to examine the patterns of social support and to identify keywords and their associations through text mining. In addition, through intervention, changes in the levels of social-cognitive factors among KC women were confirmed through statistical analysis. To achieve this, SNS messages exchanged among group members for 12 weeks were characterized using text analysis, text mining, and statistical methods. It is worth mentioning that, to our knowledge, this is the first attempted analysis of SNS text messages to examine patterns of social support.

Strengths

This study has a number of strengths. First, the result of the mixed methods analysis demonstrated that the most common type of social support conveyed through SNS messages was network support, followed by emotional support, information support, and esteem support. This finding is incongruent with other studies that have examined the influence of social support through an SNS. In a study of an online community for breast cancer survivors, the most common social support identified

was information support [22]. In another intervention designed to improve smoking cessation efforts, information and emotional support were found to be the most frequent supports [37]. Network support includes not only stories from the everyday lives of the health-related online community's participants, such as chat, humor, and teasing (eg, birthday wishes and vacation plans), but also stories that are not related to health problems [22,26]. KC women often interacted through the SNS because they worked as house caretakers during the week throughout the intervention. Since KC women are often in similar situations and environments, it is thought that there is a high frequency of network support interactions like real-life-related topics between participants [38]. Thus, to approach migrant populations within this environment, network support can be hypothesized to impact health promotion [39].

Second, we found that levels of social-cognitive factors after a 12-week intervention were significantly increased in groups that had a high frequency of network support. In a previous study that analyzed dialogue between Canadian Indigenous women who participated in internet chat rooms [40], those who received social support through SNS activities were encouraged to enhance their confidence and self-efficacy for health promotion. Also, network support in an online community may help to solidify the social network and increase a sense of community among group members [28]. In another study that analyzed conversations on social media sites like Facebook, text messages related to network support derived more reciprocal responses among participants, and that study suggested that (1) network support involves communication with many other people with similar experiences and (2) there may be evidence that network support increases a sense of belonging to a social group [21,41]. This explanation is applicable to the findings of this study because the group with the most frequent network support experienced a significant increase in sense of community after 12 weeks. As in previous studies, it was shown to be necessary to increase social-cognitive factors for exercise adherence; social support facilitated an increase in social-cognitive factors [12,13]. Therefore, strengthening network support through an SNS in an online community may be useful for increasing physical activity.

Third, this study was the first to attempt text mining to analyze text messages shared via SNS among KC women. Quantitative analysis has limitations when trying to explore the real-life social network of participants. Through text mining, this study extracted interests, themes, and relationships of words using scientific evidence. In a study of alcohol and marijuana use, online text message analysis using text mining was shown to improve understanding of the participants' impact on health-related behaviors and the role of social networks [42]. Also, by using text mining to analyze patterns of participants' chats about electronic cigarette cessation, a study attempted to find factors of interest in developing strategies for promoting future behavioral change [43]. Specifically, co-occurrence word analysis confirmed the development process and structural relationships of scientific knowledge by extracting research topics and discovering the relationships among them [44,45].

After identifying the most common keywords (from #1 to #20) using a keyword-extraction algorithm, it was noted that KC middle-aged women frequently talked about exercise. This study confirmed that the use of an SNS encouraged participants to keep exercising because exercise was perceived as routine or a habit rather than as an extra burden [46]. Although it has been suggested that more evidence is needed to determine the actual usability of an SNS for changing behavior and increasing engagement [47], this study suggests that network support through an SNS may be an effective strategy for exercise adherence. Also, KC middle-aged women mainly talked about the contents of offline cultural classes addressing acculturation. Even if programs for cultural adaptation were provided to immigrants, low participation rates and high attrition rates would be likely be obstacles, as has been noted [46,48]. Although immigrants lack social networks and social support due to language or cultural differences, it was found that SNS-based group intervention can be a useful tool for enhancing the intercultural interaction among immigrants [48].

Moreover, co-occurrence analysis revealed which keywords were related to awareness of the promotion of physical activity. Specifically, *We*, *Exercise*, and *Happiness* were the most closely connected terms; the use of *We* in the chat group indicated that participants often thought collectively, and *We – Exercise* means that they encouraged one another to exercise and be healthy and perceived each other as a community. Another mobile phone–based intervention showed consistent findings through a walking group or buddies to enhance social support for exercise [49]. Moreover, the words *We* and *Happiness* were highly related, implying that SNS-mediated social support created a network of support among group members, leading to happiness in life as well as a form of exercise adherence. Therefore, an SNS-based group intervention is expected to be effective in forming network support and a sense of community and to promote health through exercise adherence for groups with limited access to health care. Studies that determine if network support can predict health outcomes such as happiness are apparently warranted.

Limitations

Text analysis of SNS messages has several limitations; therefore, care is needed when interpreting the implications of these findings. First, some messages contain two or more types of social support. Since this study set the criteria for dialogue as the *send* button, within one dialogue phrase there can be examples of mixed social support. For example, a greeting, which corresponds with *emotional support*, and teaching, which corresponds with *information support*, can coexist in one dialogue message. Second, an additional type of social support might be needed. In this study, we identified a pattern for asking questions and finding answers in the SNS. Although this study classifies this conversation type as *network support*, it may also be further broken down into *seeking support* and *providing support* [22]. Third, there was a lack of data saturation. A total of 259 SNS text messages collected over 12 weeks was insufficient for text mining. Although the KoALA program, which is suitable for the Korean language, was used, 79 conversations included emoticons or URL video links and, thus, were excluded from text mining. As emoticons enrich the information provided on SNSs and strengthen social connections, further research to analyze emoticons is required,

which will enhance usability [50-52]. Lastly, participation patterns in chats differed by group, and although positive effects through the SNS may occur, the potential for negative consequences (eg, loneliness and social comparison) cannot be excluded [47]. Therefore, further research must examine the potential negative influences of participation in SNSs and must suggest strategies to maximize the likelihood of positive outcomes following the use of an SNS.

Conclusions

SNSs have been found to be essential communication tools for KC migrant women in Korea. The findings of this study improve our understanding of the potential impact of SNSs on improving social-cognitive factors related to exercise and promoting healthy behaviors in migrants. Specifically, the KC migrant women who participated in the culturally adaptive walking intervention described herein received network support from group members through the SNS. The group network support appeared to be more effective at increasing the levels of social-cognitive factors, which are known to be facilitators of healthy behaviors and acculturation, compared with other types of social support. Therefore, interventions designed to enhance network social support through SNSs may be effective strategies for improving health-related behaviors and cultural adaptation for migrant populations.

Acknowledgments

This work was supported by a grant from the National Research Foundation (NRF) of Korea, funded by the Korean government (Ministry of Science and ICT) (NRF-2017R1A2B4008671) and the Ministry of Education (NRF-2020R1I1A2069894). We would like to thank the Korea Support Center for Foreign Workers for their cooperation.

Conflicts of Interest

None declared.

References

1. Shepherd A, Sanders C, Doyle M, Shaw J. Using social media for support and feedback by mental health service users: Thematic analysis of a twitter conversation. BMC Psychiatry 2015 Feb 19;15:29 [FREE Full text] [doi: 10.1186/s12888-015-0408-y] [Medline: 25881089]
2. Xu Y, Burleson BR. Effects of sex, culture, and support type on perceptions of spousal social support: An assessment of the "support gap" hypothesis in early marriage. Hum Commun Res 2001 Oct;27(4):535-566. [doi: 10.1111/j.1468-2958.2001.tb00792.x]
3. Erfani SS, Blount Y, Abedin B. The influence of health-specific social network site use on the psychological well-being of cancer-affected people. J Am Med Inform Assoc 2016 May;23(3):467-476. [doi: 10.1093/jamia/ocv170] [Medline: 26911816]
4. Park BK, Calamaro C. A systematic review of social networking sites: Innovative platforms for health research targeting adolescents and young adults. J Nurs Scholarsh 2013 Sep;45(3):256-264. [doi: 10.1111/jnu.12032] [Medline: 23676115]
5. Korea Immigration Service Statistics 2017. Daejeon, Korea: Statistics Korea; 2017. URL: https://tinyurl.com/y6mzzqzs [accessed 2020-10-01]
6. Piao Z. A Study on the Health Status Among Middle-Aged Korean-Chinese Women Workers in Korea [master's thesis]. Seoul, Korea: Seoul National University; 2013. URL: http://s-space.snu.ac.kr/bitstream/10371/124794/1/000000013744.pdf [accessed 2020-04-07]
7. Seifi B, Ghanizadeh G, Seyedin H. Disaster health literacy of middle-aged women. J Menopausal Med 2018 Dec;24(3):150-154 [FREE Full text] [doi: 10.6118/jmm.2018.24.3.150] [Medline: 30671406]
8. Chee W, Kim S, Chu T, Tsai H, Ji X, Zhang J, et al. Practical issues in developing a culturally tailored physical activity promotion program for Chinese and Korean American midlife women: A pilot study. J Med Internet Res 2016 Nov 21;18(11):e303 [FREE Full text] [doi: 10.2196/jmir.6454] [Medline: 27872035]
9. Nerhus M, Berg AO, Haram M, Kvitland LR, Andreassen OA, Melle I. Migrant background and ethnic minority status as predictors for duration of untreated psychosis. Early Interv Psychiatry 2015 Feb;9(1):61-65. [doi: 10.1111/eip.12106] [Medline: 24225002]
10. Lee H, Chae D, Lee K, Lee M. Experiences of middle-aged Korean-Chinese female migrant workers in Korea: With focus on risk factors in work-related musculoskeletal diseases. J Korean Acad Community Health Nurs 2013;24(2):185. [doi: 10.12799/jkachn.2013.24.2.185]
11. Lee H, Wilbur J, Chae D, Lee K, Lee M. Barriers to performing stretching exercises among Korean-Chinese female migrant workers in Korea. Public Health Nurs 2015;32(2):112-121. [doi: 10.1111/phn.12105] [Medline: 24447032]
12. Cooper LB, Mentz RJ, Sun J, Schulte PJ, Fleg JL, Cooper LS, et al. Psychosocial factors, exercise adherence, and outcomes in heart failure patients: Insights from heart failure: A controlled trial investigating outcomes of exercise training (HF-ACTION). Circ Heart Fail 2015 Nov;8(6):1044-1051 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.115.002327] [Medline: 26578668]

13. Bender MS, Cooper BA, Park LG, Padash S, Arai S. A feasible and efficacious mobile-phone based lifestyle intervention for Filipino Americans with type 2 diabetes: Randomized controlled trial. JMIR Diabetes 2017 Dec 12;2(2):e30 [FREE Full text] [doi: 10.2196/diabetes.8156] [Medline: 30291068]
14. Stragier J, Mechant P, De Marez L, Cardon G. Computer-mediated social support for physical activity: A content analysis. Health Educ Behav 2018 Feb;45(1):124-131. [doi: 10.1177/1090198117703055] [Medline: 28415860]
15. Park A, Conway M, Chen AT. Examining thematic similarity, difference, and membership in three online mental health communities from Reddit: A text mining and visualization approach. Comput Human Behav 2018 Jan;78:98-112 [FREE Full text] [doi: 10.1016/j.chb.2017.09.001] [Medline: 29456286]
16. Chair SY, Wong KB, Tang JY, Wang Q, Cheng HY. Social support as a predictor of diet and exercise self-efficacy in patients with coronary artery disease. Contemp Nurse 2015;51(2-3):188-199. [doi: 10.1080/10376178.2016.1171726] [Medline: 27030520]
17. Lee D, Young SJ. Investigating the effects of behavioral change, social support, and self-efficacy in physical activity in a collectivistic culture: Application of stages of motivational readiness for change in Korean young adults. Prev Med Rep 2018 Jun;10:204-209 [FREE Full text] [doi: 10.1016/j.pmedr.2018.03.001] [Medline: 29868369]
18. Tang F, Chi I, Dong X. The relationship of social engagement and social support with sense of community. J Gerontol A Biol Sci Med Sci 2017 Jul 01;72(suppl_1):S102-S107 [FREE Full text] [doi: 10.1093/gerona/glw187] [Medline: 28575257]
19. House JS. Work Stress and Social Support. Boston, MA: Addison-Wesley Pub Co; 1983.
20. Cohen S, Syme S, editors. Social Support and Health. Cambridge, MA: Academic Press; 1985.
21. Cutrona CE, Suhr JA. Controllability of stressful events and satisfaction with spouse support behaviors. Communic Res 2016 Jun 30;19(2):154-174. [doi: 10.1177/009365092019002002]
22. Wang X, Zhao K, Street N. Analyzing and predicting user participations in online health communities: A social support perspective. J Med Internet Res 2017 Apr 24;19(4):e130 [FREE Full text] [doi: 10.2196/jmir.6834] [Medline: 28438725]
23. Fujitani T, Ohara K, Kouda K, Mase T, Miyawaki C, Momoi K, et al. Association of social support with gratitude and sense of coherence in Japanese young women: A cross-sectional study. Psychol Res Behav Manag 2017;10:195-200 [FREE Full text] [doi: 10.2147/PRBM.S137374] [Medline: 28721109]
24. Lin N, Simeone R, Ensel W, Kuo W. Social support, stressful life events, and illness: A model and an empirical test. J Health Soc Behav 1979 Jun;20(2):108-119. [Medline: 479524]
25. O'Reilly CA, Chatman J. Organizational commitment and psychological attachment: The effects of compliance, identification, and internalization on prosocial behavior. J Appl Psychol 1986;71(3):492-499. [doi: 10.1037/0021-9010.71.3.492]
26. Braithwaite DO, Waldron VR, Finn J. Communication of social support in computer-mediated groups for people with disabilities. Health Commun 1999;11(2):123-151. [doi: 10.1207/s15327027hc1102_2] [Medline: 16370973]
27. Cobb S. Presidential Address-1976. Social support as a moderator of life stress. Psychosom Med 1976;38(5):300-314. [doi: 10.1097/00006842-197609000-00003] [Medline: 981490]
28. Wang X, Zuo Z, Zhao K. The evolution and diffusion of user roles in online health communities: A social support perspective. In: Proceedings of the 3rd International Conference on Healthcare Informatics (ICHI 2015). 2015 Presented at: 3rd International Conference on Healthcare Informatics (ICHI 2015); October 21-23, 2015; Dallas, TX p. 21-23. [doi: 10.1109/ichi.2015.12]
29. Kim Y, Lee H, Lee MK, Lee H, Jang H. Development of a living lab for a mobile-based health program for Korean-Chinese working women in South Korea: Mixed methods study. JMIR Mhealth Uhealth 2020 Jan 08;8(1):e15359 [FREE Full text] [doi: 10.2196/15359] [Medline: 31913134]
30. Jeon BJ, Choi YJ, Kim HW. Application development for text mining: KoALA. Inf Syst Rev 2019 May 31;21(1):117-137. [doi: 10.14329/isr.2019.21.2.117]
31. Peterson NA, Speer PW, McMillan DW. Validation of a Brief Sense of Community Scale: Confirmation of the principal theory of sense of community. J Community Psychol 2007 Jan;36(1):61-73. [doi: 10.1002/jcop.20217]
32. Oh J, Kim C, Ha M. A study on effect of adolescents' leisure satisfaction on self-esteem and sense of community: Focused on vocational high school students. Korea J Hosp Res 2012;26(2):23-40 [FREE Full text]
33. McAuley E, Courneya KS. Self-efficacy relationships with affective and exertion responses to exercise. J Appl Soc Psychol 1992 Feb;22(4):312-326. [doi: 10.1111/j.1559-1816.1992.tb01542.x]
34. Choi M, Kim H. Nutrition knowledge, dietary self-efficacy and eating habits according to student's stage of regular breakfast or exercise [article in Korean]. Korean J Community Nutr 2008 Sep 29;13(5):653-662 [FREE Full text]
35. Sallis JF, Grossman RM, Pinski RB, Patterson TL, Nader PR. The development of scales to measure social support for diet and exercise behaviors. Prev Med 1987 Nov;16(6):825-836. [doi: 10.1016/0091-7435(87)90022-3] [Medline: 3432232]
36. Choi J. Construction of leisure physical activity model of middle-aged women in urban area. Korean J Adult Nurs 2008 Aug 19;20(4):626-640 [FREE Full text]
37. Granado-Font E, Ferré-Grau C, Rey-Reñones C, Pons-Vigués M, Pujol Ribera E, Berenguera A, et al. Coping strategies and social support in a mobile phone chat app designed to support smoking cessation: Qualitative analysis. JMIR Mhealth Uhealth 2018 Dec 20;6(12):e11071 [FREE Full text] [doi: 10.2196/11071] [Medline: 30573445]

38. Coulson NS, Buchanan H, Aubeeluck A. Social support in cyberspace: A content analysis of communication within a Huntington's disease online support group. Patient Educ Couns 2007 Oct;68(2):173-178. [doi: 10.1016/j.pec.2007.06.002] [Medline: 17629440]
39. Frohlich DO. The social support model for people with chronic health conditions: A proposal for future research. Soc Theory Health 2014 Apr 17;12(2):218-234. [doi: 10.1057/sth.2014.3]
40. Wright K. Social networks, interpersonal social support, and health outcomes: A health communication perspective. Front Commun 2016 Oct 14;1:1-6. [doi: 10.3389/fcomm.2016.00010]
41. Andalibi N, Haimson OL, Choudhury MD, Forte A. Social support, reciprocity, and anonymity in responses to sexual abuse disclosures on social media. ACM Trans Comput Hum Interact 2018 Oct 17;25(5):1-35. [doi: 10.1145/3234942]
42. Pizzicato LN, Barbour RC, Kershaw T. Evaluating alcohol and marijuana use among emerging adult males via analysis of text messages. Subst Use Misuse 2019;54(9):1547-1557. [doi: 10.1080/10826084.2019.1594902] [Medline: 31023137]
43. Chen AT, Zhu S, Conway M. What online communities can tell us about electronic cigarettes and hookah use: A study using text mining and visualization techniques. J Med Internet Res 2015 Sep 29;17(9):e220 [FREE Full text] [doi: 10.2196/jmir.4517] [Medline: 26420469]
44. He Q. Knowledge discovery through co-word analysis. Libr Trends 1999;48(1):133-159.
45. Small H. Visualizing science by citation mapping. J Am Soc Inf Sci 1999 Jun 22;50(9):799-813. [doi: 10.1002/(sici)1097-4571(1999)50:9<799::aid-asi9>3.0.co;2-g]
46. Laranjo L, Neves AL, Costa A, Ribeiro RT, Couto L, Sá AB. Facilitators, barriers and expectations in the self-management of type 2 diabetes--A qualitative study from Portugal. Eur J Gen Pract 2015 Jun;21(2):103-110. [doi: 10.3109/13814788.2014.1000855] [Medline: 25698085]
47. Balatsoukas P, Kennedy CM, Buchan I, Powell J, Ainsworth J. The role of social network technologies in online health promotion: A narrative review of theoretical and empirical factors influencing intervention effectiveness. J Med Internet Res 2015 Jun 11;17(6):e141 [FREE Full text] [doi: 10.2196/jmir.3662] [Medline: 26068087]
48. Erdem B. US-Turkey relations in the light of Turkey's freedom of press. Glob Media J 2018;16(30):1-7 [FREE Full text]
49. Lindgren T, Hooper J, Fukuoka Y. Perceptions and experiences of women participating in a digital technology-based physical activity intervention (the mPED Trial): Qualitative study. JMIR Public Health Surveill 2019 Dec 20;5(4):e13570 [FREE Full text] [doi: 10.2196/13570] [Medline: 31859677]
50. Cui M, Jin Y, Kwon O. A method of analyzing sentiment polarity of multilingual social media: A case of Korean-Chinese languages [article in Korean]. J Intell Inf Syst 2016 Sep 30;22(3):91-111 [FREE Full text] [doi: 10.13088/jiis.2016.22.3.091]
51. Edo-Osagie O, Smith G, Lake I, Edeghere O, De La Iglesia B. Twitter mining using semi-supervised classification for relevance filtering in syndromic surveillance. PLoS One 2019;14(7):e0210689 [FREE Full text] [doi: 10.1371/journal.pone.0210689] [Medline: 31318885]
52. Hsieh SH, Tseng TH. Playfulness in mobile instant messaging: Examining the influence of emoticons and text messaging on social interaction. Comput Human Behav 2017 Apr;69:405-414. [doi: 10.1016/j.chb.2016.12.052]

Abbreviations

KC: Korean-Chinese
KoALA: Korean Natural Language Application
NRF: National Research Foundation
SNS: social networking service

Edited by L Buis; submitted 07.04.20; peer-reviewed by C Rey-Reñones, J Gregg; comments to author 12.06.20; revised version received 23.07.20; accepted 11.08.20; published 05.11.20.

Please cite as:
Lee H, Lee H, Kim Y, Kim S, Lee YM
Network Support Using Social Networking Services to Increase Exercise Adherence Among Korean-Chinese Middle-Aged Migrant Women: Mixed Methods Study
JMIR Mhealth Uhealth 2020;8(11):e19159
URL: https://mhealth.jmir.org/2020/11/e19159
doi:10.2196/19159
PMID:33151155

Original Paper

Gamified Text Messaging Contingent on Device-Measured Steps: Randomized Feasibility Study of a Physical Activity Intervention for Cancer Survivors

Michael C Robertson[1,2], MPH; Elizabeth J Lyons[3], PhD, MPH; Yue Liao[1,4], PhD, MPH; Miranda L Baum[1], MA; Karen M Basen-Engquist[1], PhD, MPH

[1]Department of Behavioral Science, University of Texas MD Anderson Cancer Center, Houston, TX, United States

[2]Health Promotion & Behavioral Sciences, University of Texas School of Public Health, Houston, TX, United States

[3]Department of Nutrition and Metabolism, School of Health Professions, The University of Texas Medical Branch at Galveston, Galveston, TX, United States

[4]Department of Kinesiology, College of Nursing and Health Innovation, University of Texas at Arlington, Arlington, TX, United States

Corresponding Author:
Karen M Basen-Engquist, PhD, MPH
Department of Behavioral Science
University of Texas MD Anderson Cancer Center
1515 Holcombe Blvd
Houston, TX, 77030
United States
Phone: 1 713 745 3123
Email: kbasenen@mdanderson.org

Abstract

Background: Physical activity can confer diverse benefits on cancer survivors. Unfortunately, many cancer survivors are not sufficiently active. The efficacy of physical activity interventions for this population may be increased by grounding them in Self-Determination Theory (SDT). Combining game design elements with wearable technologies may be a useful and scalable approach to targeting SDT constructs to promote cancer survivors' physical activity.

Objective: The primary aim of this study is to evaluate the feasibility and acceptability of *Steps2Health*, a physical activity intervention for cancer survivors. It also aims to investigate the effects of the intervention on motivation, physical activity, and step count.

Methods: We randomized 78 insufficiently active cancer survivors to an experimental or comparison group. All participants received a physical activity tracker. The experimental group participants also received a set sequence of multimedia messaging service messages that were triggered in real time by meeting predetermined cumulative step count totals. Messages presented information about a virtual journey and included photographs and vivid descriptions of locations to increase autonomous motivation. Additional messages targeted perceptions of *relatedness* (eg, role modeling) and *competence* (eg, facilitating mastery experiences). We administered pre- and postintervention surveys and conducted 15 individual interviews to evaluate the intervention. We performed directed content analysis of qualitative data and conducted mixed effects linear modeling to investigate participants' changes in motivation, self-reported physical activity, and device-measured step counts.

Results: There was minimal loss to follow-up (3/78, 4%), the device wear rate was high (2548/3044, 83.71% of days), and technical problems with messaging based on real-time step counts were limited. Our qualitative data analysis revealed 3 overarching themes: *accessibility*, *autonomous motivation*, and *relatedness*. Participants successfully navigated the technological aspects and game design elements of the intervention. Participants found messages targeting *autonomous motivation* and *competence or self-efficacy* to be enjoyable and compelling, but one feasibility criterion for participant engagement (response rate to text messages) was not met. Messages targeting *relatedness* were less highly rated than the messages targeting *autonomous motivation* and *competence or self-efficacy*. During the intervention, both groups increased their motivation for physical activity (B=0.16; 95% CI 0.01 to 0.30; P=.04; d=0.49), and assignment to the experimental group was associated with increased self-reported leisure activity score (B=10.78; 95% CI 3.54 to 18.02; P=.005; d=0.64). The experimental group had greater increases in daily step counts over time (B=322.08; 95% CI 54.01 to 590.15; P=.02; d=0.28).

Conclusions: This study supports the feasibility of using real-time game design elements to target SDT constructs and increase cancer survivors' physical activity. Overall, our findings support the acceptability of the *Steps2Health* intervention, but fostering active participant engagement and targeting *relatedness* may present additional challenges. *Steps2Health* may help cancer survivors increase their physical activity levels.

(JMIR Mhealth Uhealth 2020;8(11):e18364) doi:10.2196/18364

KEYWORDS

cancer survivors; physical activity; motivation; self-control; mobile health; mobile phone; technology

Introduction

Background

Physical activity is generally safe and a health protective factor for cancer survivors. It is associated with a lower risk of all-cause mortality in this population, and for survivors of some types of cancer, it is associated with lower risks of recurrence and cancer-related mortality [1]. Increasing moderate-to-vigorous physical activity may ameliorate symptoms that cancer survivors commonly report, including fatigue, pain, anxiety, decreased physical functioning, and cancer-related cognitive impairment [2]. Unfortunately, the majority of cancer survivors do not meet the nationally recommended aerobic physical activity guidelines for adults [3], which call for engaging in 150 min of moderate-intensity aerobic physical activity or 75 min of vigorous-intensity aerobic physical activity (or some equivalent combination of both) per week. An analysis of data from adult cancer survivors responding to the 2014 National Health Interview Survey indicated that under 45% of adult survivors aged 45-64 years met the guidelines for aerobic physical activity, and this percentage was approximately 35% in survivors 65 years and older [4]. Studies using accelerometers to objectively measure physical activity guideline adherence indicate that this percentage may be between 4% and 13% [5,6]. Cancer survivors may encounter barriers that impede physical activity guideline adherence faced by the general population (eg, competing demands for time) and additional barriers attributable to cancer and its treatment. These can include decreased physical functioning, fatigue, and pain [7,8].

Electronically delivered behavioral interventions have been shown to be effective in promoting physical activity in cancer survivors [9]. Such interventions readily lend themselves to widespread dissemination and increasingly feature mobile and wearable computing technologies (mobile health [mHealth]) that can provide timely feedback on behavior. Although behavioral interventions incorporating wearable consumer technologies have been used to initiate physical activity among cancer survivors [10], evidence of the long-term effectiveness of these interventions is lacking. Furthermore, most mHealth programs have high discontinuation rates (eg, 25% of users abandon mobile apps after just a single use) [11]. Many mHealth programs are centered on facilitating self-regulatory processes (eg, goal setting, self-monitoring) but are not necessarily designed to affect participants' *motivation* for physical activity. This is an important distinction, particularly because cancer survivors may face barriers that can diminish their motivation to engage in physical activity (eg, decreased physical functioning, fatigue, pain) [7,8]. Increasing cancer survivors' motivation for physical activity may facilitate sustained engagement with mHealth programs and improve long-term behavioral adherence to recommended guidelines [12].

Self-Determination Theory (SDT) provides a framework for understanding the role of quality of one's motivation in long-term adherence to health-related lifestyle behaviors [13]. It posits that motivation can be conceptualized as existing on a continuum that ranges from fully external or controlled motivation to fully internal or autonomous motivation and that increasing the latter tends to yield longer lasting behavioral change than the former [13]. SDT also holds that *autonomous regulation* is determined in large part by the satisfaction of an individual's core psychological needs (*autonomy, relatedness*, and *competence*) [13]. This theoretical approach can be useful in predicting and influencing physical activity in the general population [14], and targeting SDT constructs may be similarly beneficial for promoting physical activity in cancer survivors [15]. Indeed, SDT is increasingly being used to inform interventions in this population [16,17].

One approach that has been used to effectively target SDT constructs in the general population is gamification, the application of game design elements to nongame contexts [18-20]. Thus, we drew from the games-for-health literature to develop *Steps2Health*, an mHealth intervention that targets SDT constructs to promote cancer survivors' physical activity. As only few models for systematic planning of gamification interventions exist [21], we adapted the Behaviour Change Wheel model for this purpose [22]. This model typically asks planners to match intervention functions (eg, persuasion) to theoretical constructs (eg, reflective motivation). Our adaptation included playful experiences taken from the Playful Experiences Framework [23] as potential intervention functions. We chose to focus on playful experiences as broad methods of intervention (rather than game mechanics as specific behavior change techniques) to emphasize the autonomy-supportive, playful aspects of games, as has been recommended [21,24]. Specifically, we focused on *the playful experiences of discovery, exploration, and humor to target intrinsic regulation*. In addition, we included techniques from Motivational Interviewing and Acceptance and Commitment Therapy to complement the game messages and target *integrated regulation* [25,26].

Objectives

The main purpose of this study is to assess the feasibility of *Steps2Health* and the participating cancer survivors' satisfaction with it. Its secondary objective is to assess the effects of *Steps2Health* on participants' autonomous regulation,

self-reported physical activity, and device-measured physical activity (ie, step counts).

Methods

Recruitment

We identified potential participants through various means, including health fairs, conferences, and other local events in South Texas; social media (eg, Facebook, Twitter); and our institutional website and publications. We contacted interested individuals via email and telephone and conducted a formal screening process via telephone. We engaged in a verbal informed consent process with all eligible individuals between September 2018 and February 2019. All research protocols were approved by the University of Texas MD Anderson Cancer Center's Institutional Review Board (protocol 2018-0239). Participants were adult cancer survivors who had completed primary cancer treatment for at least three months previously, owned a smartphone, and were willing to receive text messages and complete web-based surveys. Eligible participants did not meet the recommended physical activity levels [27] at screening as determined by verbal administration of the modified Godin Leisure-Time Exercise Questionnaire [28].

Study Design

We conducted a randomized controlled pilot trial. As the intervention duration was contingent on participants' cumulative step count (ie, participants who registered more daily steps progressed though the intervention more quickly), we assigned participants to the experimental and comparison groups in pairs. We recruited cohorts of 6 participants (to facilitate the logistics of study operations) and randomly assigned pairs within each cohort and group assignment within each pair (ie, for each group of 6 participants, we randomly assigned each participant to 1 of 3 pairs, and for each pair, we randomly assigned one participant to the experimental group and the other to the comparison group). The study staff conducted this restricted randomization procedure using Research Electronic Data Capture (Vanderbilt University). The intervention duration for both participants in each pair was determined by the experimental group participant's time to reach a predetermined cumulative step count (see the Intervention section). We administered surveys before and after the intervention. For each pair, we continued to record participants' step counts for 4 weeks after the experimental group participant had completed the intervention. The participants assigned to the control arm completed baseline and follow-up assessments in the same timeframe as their paired counterpart (ie, the Intervention section details the study duration for each pair determined by the activity level of the experimental group participant).

We conducted individual, semistructured interviews with 15 experimental group participants at the end of their study participation. Questions were centered on obtaining feedback on the feasibility and acceptability of the intervention and insights into how it may be improved (Multimedia Appendix 1). We interviewed some participants who completed the intervention most quickly and some who completed it least quickly. We interviewed additional participants at the discretion of the principal investigator (eg, to ensure that we interviewed some men). Each interview lasted for 30 to 60 min, and all interviews were conducted by the first author (MR). The study staff (MB) took detailed notes in all interviews. Immediately after the interview, the first author reviewed and contributed to these notes. We performed qualitative data collection until there was consensus among the research team that the point of data saturation had been reached and the individual interviews were not producing novel content any longer.

Intervention

All participants engaged in this study remotely. We synchronized wrist-worn Fitbit Alta devices to Fitabase (Small Steps Laps), a web-based data management platform for wearable devices, and then mailed each participant a device. We instructed participants to wear the device during waking hours. We included instructions and links to the Fitbit content detailing how to set up the Fitbit device to synchronize automatically. All participants had access to the Fitbit website and app. In addition, the participants in the experimental group received *Steps2Health* multimedia messaging service (MMS) messages that were designed to target SDT constructs. These messages, which were developed by the research team, presented information about a virtual journey through Japan's Inland Sea region and were triggered by step counts in real time. Participants' progress on the 166,000-step (approximately 83 miles) virtual journey was determined by cumulative step counts measured by the Fitbit devices. All experimental group participants received the same series of text messages. The duration of the *Steps2Health* intervention was determined by how long it took participants to register 166,000 steps on their device (ie, if participants were more active, they received the messages more frequently). We chose this distance because we anticipated that it would take most participants about 1 month to complete the journey. We designed the *Steps2Health* intervention to have a variable duration contingent on participant step count to facilitate a sense of autonomy for the user and verisimilitude for the gamified intervention's premise of undertaking an actual journey in real time. There is only one journey featured in this pilot study, but ultimately, *Steps2Health* may feature a variety of journey options that users can engage in sequentially for a more continuous experience.

The *Steps2Health* intervention included 54 total messaging blocks designed to target *autonomous regulation*, *autonomy*, *relatedness*, and *competence* (Multimedia Appendix 2). Messages were informed by insight from previous qualitative research in cancer survivors and developed in a consensus-building process among the research team, which has expertise in health behavior change theory and physical activity promotion for cancer survivors [29]. All messaging blocks contained text, and some additionally contained images or hyperlinks (Table 1). Each block included 1 to 3 messages. Messages targeting *autonomous regulation* presented photographs and vivid descriptions of destinations along a geographically accurate virtual tour of the region. These messages presented playful experiences of exploration, discovery, and humor and encouraged participants to identify value-based life goals linked to physical activity (Table 1). Some messages contained hyperlinks to various resources for healthy living (eg, videos demonstrating muscle strengthening

exercises, healthy recipes local to the current virtual location, stress-reduction techniques, etc). We designed the *Steps2Health* intervention to most heavily target this construct (devoting 21 of 54 messaging blocks to it), given its emphasis on SDT and its central role in the gamification elements of the intervention. Participants were also given the choice to participate in additional mini journeys (13 messaging blocks). These optional messages were similar to those targeting *autonomous regulation*. To target the SDT construct of *relatedness*, we included messages in 10 messaging blocks from Ruby, an ovarian cancer survivor (this was a fictitious character, and participants were made aware of this). This character was written to be a positive role model. Her messages provided encouragement and prompted participants to reflect on questions that were derived from motivational interviewing principles to elicit positive *change talk* [26]. Finally, 10 messaging blocks included messages adapted from previous studies to increase participants' *competence or self-efficacy* for increasing physical activity.

Table 1. Example messages targeting Self-Determination Theory constructs.

Step count	Example messages	Image	Construct targeted
1	• Welcome to Steps2Health! • Please save this number in your phone as Steps2Health, and be sure that your Fitbit is up set to sync automatically. • Bridges serve as major checkpoints for this 83 mile island-hopping trek through beautiful Japanese islands. Keep your step count high to maximize your progress!	Starting message image	Autonomous motivation
8000	• RUBY: Hello! My name is Ruby. I am an ovarian cancer survivor and have already completed this journey. • I wanted to get strong to keep up with my grandson. • Is there a goal you'd like to work toward? • Would you share it with me in a text? If not, just text 0.	Relatedness example message image	Relatedness
45,000	• Would you like to take a quick trip to Onomichi, the "Town of Hills and Cats" today? You'll get some extra photos of high points of Onomichi. • Reply YES or NO	N/A[a]	Autonomy or autonomous motivation
57,300	• You have made it to the beautiful Kosanji temple. It was built in 1936 by a wealthy industrialist in honor of his mother! • It is written in a famous haiku: • The mothers of the world are as the Goddess of Mercy.	Autonomous motivation example message image	Autonomous motivation
84,000	HEALTH TIP: Living through cancer can be stressful, but you can manage the stress. Even a 10-minute time-out can help by taking time to move and breathe.	N/A	Competence or self-efficacy

[a]N/A: not applicable.

We used Fitabase to gather participants' data in real time from the Fitbit server. We worked with Mosio to build a platform to send intervention text messages on the basis of participants' real-time step counts. This platform queried Fitabase twice an hour to determine participants' cumulative step counts. We ensured that participants enabled their devices to frequently and automatically upload current data to the Fitbit server (ie, synchronize with the server). The text messaging platform sent automatic text message reminders for participants to synchronize their devices if they had not done so in 2 days or more, and study staff sent email reminders after 5 days.

Outcomes

Overview

For the primary aim of this study, we evaluated intervention feasibility and acceptability via quantitative and qualitative methods. For the secondary aim, we evaluated changes in autonomous regulation, self-reported moderate-to-vigorous physical activity, and device-measured step count over the course of the intervention period. We administered web-based surveys to each participant at baseline and after the experimental group participants completed the *Steps2Health* intervention. The surveys contained items pertaining to participants' characteristics, the feasibility of the intervention, and participants' autonomous motivation and physical activity.

Intervention Feasibility

We evaluated feasibility by assessing whether (1) more than 75% (59/78) of all participants wore their Fitbit at least 5 days a week for at least 75% of the study period (*valid wear*); (2) less than 30% (23/78) of all participants reported technical difficulties with the Fitbit or receipt of the text messages (*technical difficulties*); (3) the experimental group participants responded to at least 80.1% (157/196) of the messages requiring a response (*participant engagement*); and (4) at least 75% (30/39) of the experimental group participants reported that they would recommend the program to friends or family (*participant satisfaction*). These criteria were derived from

metrics for assessing the feasibility of consumer-based wearable physical activity trackers used in other digital behavior change studies [30] and our intervention's theoretical orientation.

Autonomous Regulation

We administered the Behavioral Regulation for Exercise Questionnaire-3 (BREQ-3) to all participants before and after the intervention. BREQ-3 has acceptable internal consistency [31]. As SDT posits that *autonomous regulation* is composed of identified regulation, integrated regulation, and intrinsic regulation, we averaged these subscale scores to represent this construct [13].

Physical Activity

We administered the Godin Leisure-Time Exercise Questionnaire and calculated the Leisure Score Index to measure participants' moderate-to-vigorous–intensity physical activity [28]. This survey has minimal participant burden, good test-retest reliability (reliability coefficient=0.81), and convergent validity with VO_2 max (maximum rate of oxygen consumption during intense exercise) [28,32]. In addition, we evaluated participants' daily Fitbit step count trends. In consideration of the insufficiently active sample recruited in this study and to maximize the use of available data, we used 2 variables to define nonvalid wear days. We defined a nonvalid wear day as a day during which participants (1) did not wear the device for at least 10 hours (out of the full 24-hour day) according to minute-level device data and (2) registered fewer than 1500 steps according to day-level data [33-35]. Minute-level nonwear was defined as periods of 90 consecutive minutes of 0 steps with a 2-min nonzero tolerance, consistent with commonly used accelerometer protocols [36]. We used a combination of day- and minute-level data because some research indicates that the use of accelerometer protocols for minute-level data may tend to overestimate nonwear in Fitbit data [37], and Fitbit devices are programmed to automatically delete minute-level data when their batteries run low or if the devices are not regularly synchronized. Thus, we supplemented this decision rule with the 1500 steps day-level threshold used in other studies [34,35].

Data Analysis

Two analysts (MR and MB) conducted directed content analysis [38] of the field notes from the individual interviews. We created and assigned inductive and deductive codes to discrete points made by each participant. We organized these codes in an iterative process to gain insight into the perceived feasibility and acceptability of the intervention and means by which the intervention might be improved.

We calculated descriptive statistics for participant characteristics and feasibility items and performed Pearson chi-square tests to evaluate differences in descriptive characteristics between groups. We used linear mixed models to assess between-group differences in pre- to postintervention changes in *autonomous regulation* (as assessed using the BREQ-3) and physical activity (as assessed using the Godin Leisure-Time Exercise Questionnaire Leisure Score Index). These models included terms for group-by-time interactions with random coefficients for participants nested within pairs. We used a linear growth model with random intercepts and slopes to assess between-group differences in changes in Fitbit-measured step counts during the intervention [39]. As the intervention duration varied among pairs, we modeled each participant's average daily step count as a function of quartiles of the intervention period. We specified an autoregressive covariance structure because of the time series nature of the data and adjusted for the number of days of valid Fitbit wear. We used the maximum likelihood estimation for all linear mixed models. We performed a likelihood ratio test to determine if it was necessary to specify the third level in the linear growth model (pairs subsuming individuals). The results of this test did not show a statistically significant difference in model fit between the 3-level model and a 2-level model (P=.90); furthermore, the conclusions to be drawn from the results of the competing models were not substantively different. Accordingly, we present results from the more parsimonious 2-level model with observations nested in individuals. We created plots recommended by Bolger and Laurenceau [40] to analyze and present the linear growth model results. We supplemented the mixed model findings with Cohen d effect size calculations. We set the nominal α value to .05 for all analyses, which we performed in R version 3.6.1.

Results

Participants

We randomized 78 participants to either the experimental group or comparison group. Of these 78 participants, 3 (4%) were lost to follow-up and 75 (96%) completed the baseline and follow-up surveys (Figure 1). The study sample was mostly female (71/78, 91%) and relatively well educated (Table 2). The participants' mean age was 55.1 (SD 13.5) years. Most participants were breast cancer survivors (45/78, 58%) and overweight (38/75, 50%) or obese (19/75, 25%). The overall mean time since cancer diagnosis was 9.4 (SD 7.3) years. All participants in the experimental group who completed the preintervention survey also completed the 166,000-step journey.

Figure 1. The CONSORT (Consolidated Standards of Reporting Trials) diagram for Steps2Health recruitment, retention, and analysis.

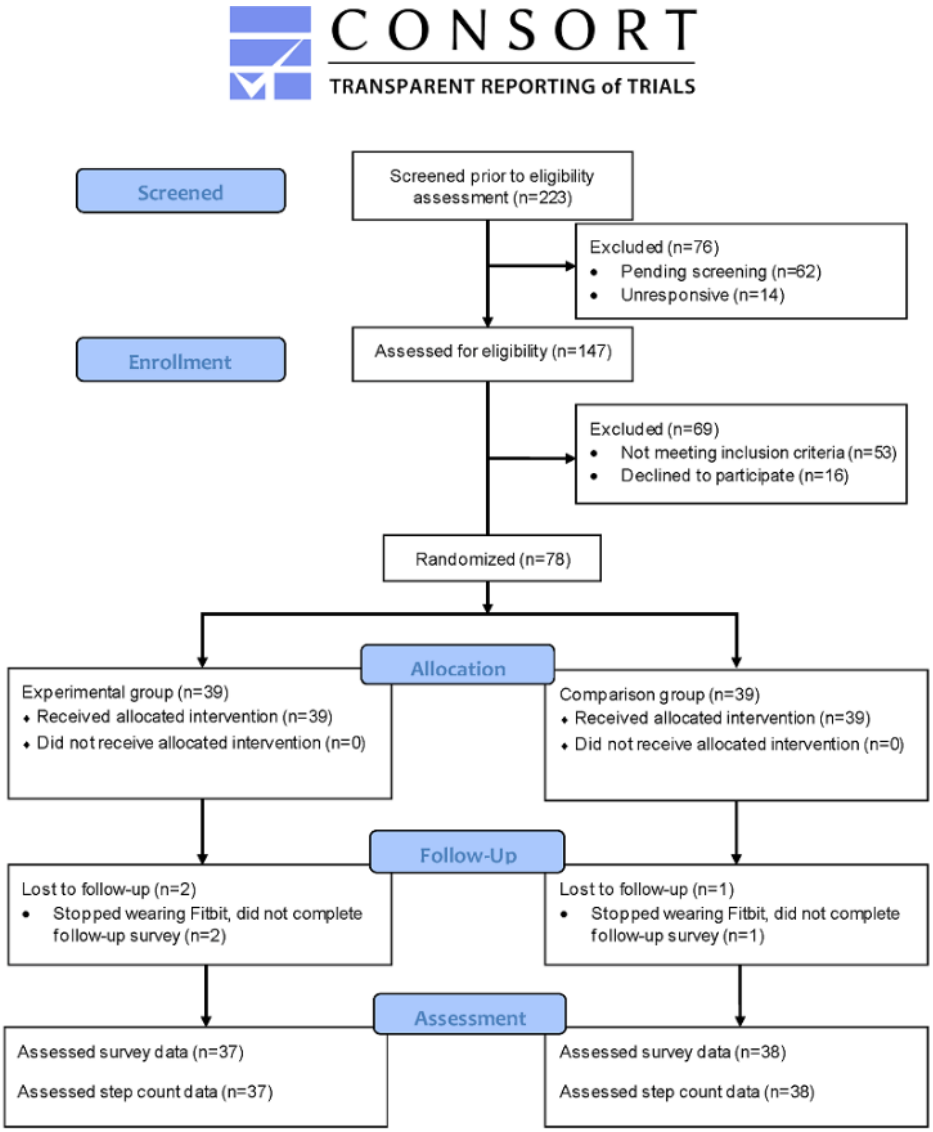


Table 2. Participant characteristics (N=78).

Characteristics	Experimental group, n (%)	Comparison group, n (%)	*P* value[a]
Age (years)			.89
18-34	2 (5)	2 (5)	
35-49	10 (26)	14 (36)	
50-64	13 (33)	12 (31)	
65-74	13 (33)	10 (26)	
≥75	1 (3)	1 (3)	
Education level			.16
<High school	0 (0)	1 (3)	
High school diploma or general educational development	5 (13)	2 (5)	
Some college	11 (28)	12 (31)	
Bachelor's degree	14 (36)	10 (26)	
Graduate school degree	9 (23)	14 (36)	
Employment status			.98
Employed full time	15 (34)	18 (42)	
Employed part time	8 (18)	6 (14)	
Not employed for pay, not seeking paid employment	2 (5)	1 (2)	
Not employed for pay, but seeking paid employment	2 (5)	1 (2)	
Retired	8 (18)	7 (16)	
Homemaker	6 (14)	7 (16)	
Student	1 (2)	1 (2)	
Volunteer	2 (5)	2 (5)	
Gender			.99
Female	36 (92)	35 (90)	
Male	3 (8)	4 (10)	
Marital status			.62
Single	9 (23)	5 (13)	
Married	23 (59)	24 (62)	
Divorced	6 (15)	9 (23)	
Widowed	1 (3)	1 (3)	
Race			.50
American Indian or Alaska native	1 (3)	0 (0)	
Asian	0 (0)	0 (0)	
Black or African American	8 (21)	6 (15)	
White	30 (77)	32 (82)	
Other	0 (0)	1 (3)	
Ethnicity[b]			.57
Hispanic	8 (21)	7 (18)	
Non-Hispanic	30 (79)	32 (82)	
Cancer diagnosis			.46
Breast	20 (51)	25 (64)	
Ovarian	4 (10)	3 (7)	

Characteristics	Experimental group, n (%)	Comparison group, n (%)	*P* value[a]
Colorectal	1 (3)	3 (7)	
Endometrial	6 (15)	2 (5)	
Prostate	2 (5)	2 (5)	
Urinary tract	0 (0)	1 (2)	
Renal or pelvic	1 (3)	0 (0)	
Brain	0 (0)	1 (2)	
Other	6 (15)	6 (15)	
Body mass index status[c]			.51
Normal	11 (30)	7 (18)	
Overweight	17 (46)	21 (55)	
Obese	9 (24)	10 (26)	

[a]Pearson chi-square test.

[b]Over 98% (77/78) responded to this item.

[c]Over 96% (75/78) responded to required items.

Device Wear

The median number of days to complete the journey for experimental group participants was 30; this ranged from 15 to 128 days (IQR 23-51). The overall day-level percentage of valid Fitbit device wear was 83.71% (2548/3044). At the week level, 79% (59/75) of participants wore their Fitbit at least 5 days a week for at least 75% of the overall study period. This exceeded the a priori target of >75%, supporting the feasibility of intervention. During the intervention period, the average percentage of nonvalid wear days was 13.5% for the experimental group and 8.8% for the comparison group. During the follow-up period, these percentages were 27.5% and 20.1%, respectively.

Intervention Feasibility

We conducted in-depth interviews with 15 participants in the experimental group (12 females and 3 males). Qualitative data analysis revealed 3 overarching themes: *accessibility*, *autonomous motivation*, and *relatedness*.

Theme 1: Accessibility

Limited Technical Difficulties

Findings from qualitative and quantitative data analyses indicated that participating cancer survivors perceived the *Steps2Health* intervention to be accessible. Regarding the intervention's technological aspects, participants perceived the intervention to be straightforward and practicable. One participant stated:

> *Not a single issue. Worked straight up no problem…it was remarkable to me that there were no issues at all.* [P60]

This sentiment is generally supported by participants' survey data. Quantitative data revealed that only 11% (8/74) of participants indicated that they had problems with using the Fitbit device, and only 14% (10/74) indicated that they had problems receiving the text messages (Figure 2). These numbers were below the a priori target of <30%, supporting intervention feasibility. Technical difficulties included 4 instances of faulty hardware; in each case, the Fitbit customer support team ultimately replaced the device. One early technical issue was that the MMS messages requiring a response were repeated. This was remedied for the third (6 persons) cohort of participants. There was one instance when all intervention messages were sent at once to a participant. With the participant's permission, we were able to get this person resituated on the journey.

Figure 2. Participant satisfaction with the Steps2Health intervention.

During interviews, several participants suggested providing a more thorough orientation process that would more fully reveal exactly how the text messaging platform worked. In general, participants who offered this suggestion reported that if they did not synchronize their devices frequently, they would receive a burst of several text message blocks in quick succession when they next synchronized their devices. Participants who experienced this said that a more thorough understanding of the intervention messaging mechanism would help them to better navigate these instances and encourage them to synchronize their devices more regularly.

Celebratory Orientation

Participants appreciated the *Steps2Health* design that celebrated participants meeting their cumulative physical activity accomplishments and rewarded them with fun and interesting content. This orientation to physical activity promotion made the *Steps2Health* intervention feel like an accessible program

that accommodates participants' fitness levels. One participant said:

> *I'm not real big on exercising and it (Steps2Health) definitely was encouraging. One thing (that was encouraging) was hitting your markers when you've made it there, and there was interesting information… Having that didn't overwhelm me and I could make the markers they were setting up for me.* [P22]

A few participants noted that they appreciated that the Fitbit device gave them credit for steps when they were walking without their smartphones, and some stated that they wished the *Steps2Health* intervention would also be able to celebrate achievements in other forms of exercise that did not necessarily increase step count (eg, swimming).

Theme 2: Autonomous Motivation

Enjoyment of Journey Messages

Messages targeting *autonomous regulation* and *competence or self-efficacy* received the highest ratings in the follow-up survey (Figure 2). These messages comprised the bulk of the *Steps2Health* intervention. Participants generally found the intervention to be enjoyable and compelling. In total, 78% (29/37) of participants reported that they would recommend the intervention to a friend or family member (Figure 2). This met the a priori target of >75%, supporting the feasibility of the intervention.

In individual interviews, 14 of the 15 participants (93%) said that they would participate in another journey if offered. Some interviewees indicated that they were motivated to be more active to see what was next on the journey. One participant said:

> *It's very easy in survivorship to have the world be overwhelming, you fall behind. You move slower than before, and it helped keep it a priority. This was very helpful—some of these messages were really pleasant, and I wanted to see what was going to come next.* [P60]

Another participant said:

> *I am coming from a year where fitness wasn't my strong point after treatment but I enjoyed the messages and was sad that it ended…I was excited to read the messages, I talked about them with my kids and it was like a family event. I showed them the milestones about the bridges we crossed. It was really neat.* [P29]

However, a few participants said that the text messages were not very enjoyable or motivating to them. One said:

> *Text messages were annoying—they came at random so I blocked it.* [P34]

Despite this, the Fitbit feedback and app functions were almost universally well received.

Linked Media Content

The use of linked media content in the messages was almost categorically cited as being a good idea, but in individual interviews, many participants said that they did not often access the links. The primary reason given for this was that the participants were too busy to give these messages enough attention when they were received. Participants appreciated that the content would remain in their text message history, but many said that they seldom went back to access them. Recommendations from participants included providing a more readily accessible repository of this content and providing messages that were more individually relevant or personally tailored (eg, messages of prescriptions of physical activity that take into account the user's general fitness level, personal preferences, and/or mobility limitations).

Motivation for Physical Activity

Although participants in general seem to have enjoyed the intervention content, results concerning whether participants felt the text messages effectively motivated them to engage in more physical activity were mixed. When asked in individual interviews, participants tended to respond that the messages did make them feel more motivated to exercise. However, qualitative data analysis suggested that participants often did not disentangle the effects of the *Steps2Health* intervention messages from those provided by the Fitbit device or app (eg, messaging from the Fitbit device that is provided if or when the user reaches 10,000 steps in a day). Furthermore, the results of the quantitative data analysis were mixed, with 51% (19/37) of the sample agreeing or strongly agreeing with the statement "Receiving the text messages motivated me to increase my physical activity," whereas 22% (8/37) disagreed or strongly disagreed with this statement and 27% (10/37) indicated they were neutral (Figure 2).

Participant Engagement

Some intervention messages explicitly requested a response, and in this study, we conceptualized participant engagement as the response patterns to these intervention messages. Messages requesting a response targeted either *autonomy* or *relatedness*. The response rate for the former was 73% (45/62) and that for the latter was 36.7% (79/215). These results did not meet the a priori target of >80% and did not support intervention feasibility. In individual interviews, participants indicated that they were sometimes unaware that they had missed these texts. This may have been partly because of the fact that these text messages could sometimes become buried by subsequent messages if participants were particularly active or if they had not synchronized their devices in a while. Participants indicated that they were often on the move when these messages were received and were unable or unwilling to attend to the messages at that time. The response rate for messages pertaining to *relatedness* was particularly low, possibly owing to lower participant satisfaction with these messages, detailed below.

Theme 3: Relatedness

Participants' satisfaction with the intervention messages targeting *relatedness* was mixed. Quantitative data indicated that these messages were less well received than the other types of messages, with 24% (9/37) of participants disagreeing or strongly disagreeing with the statement, "The text messages from Ruby were enjoyable to me" (the least well rated of the specific types of messages, as shown in Figure 2). In individual

interviews, some participants were very positive about this aspect of the program:

> *I was waiting for those messages every day… I feel like [Ruby] was my coach.* [P30]

Some reported that they did not believe that this approach had much utility in fostering feelings of *relatedness*; one participant said:

> *Ruby, I just didn't do well on that. I am not one of those people who does a lot of texting.* [P68]

Others said that they liked the idea of having a supportive role model but became disillusioned with this aspect of the intervention when they received only brief, automatic replies in response to the thoughtful messages they sent.

As noted earlier, the participant response rate to questions that explicitly requested a response was low (79/215, 36.7%). During interviews, participants offered several suggestions for improving this aspect of the intervention. One suggestion was to avoid using generic replies to participant responses to mimic an actual conversation. Another suggestion was to provide multiple character avatars for users to choose from, which might enhance relatability:

> *The idea of selecting among two or three (role models) would be awesome, and that may be all you need. At most four. We aren't going to be chatting, so don't try to make it feel like we are chatting.* [P60]

This point was emphasized by the male survivors we interviewed, who felt they did not have much in common with the female role model character. To foster feelings of *relatedness*, some participants suggested making use of text messaging or social media platforms to connect users to one another. One participant said:

> *If we could do a group text… that way maybe if someone was having a hard day we could encourage each other.* [P4]

However, this notion was met with mixed opinions. According to another participant:

> *My trust level in terms of responding to other people I do not know—I just don't do that.* [P68]

Intervention Effects

Autonomous Regulation

There was a small, statistically significant effect of time such that participants' *autonomous regulation* scores tended to increase by 8% from before to after the intervention (B=0.16; 95% CI 0.01 to 0.30; P=.04; d=0.49). The results did not support a group-by-time interaction for *autonomous regulation* (P=.59).

Godin Leisure-Time Exercise Questionnaire

Linear mixed model results indicated a statistically significant group-by-time interaction for physical activity as measured by the Godin Leisure-Time Exercise Questionnaire (B=10.78; 95% CI 3.54 to 18.02; P=.005). Assignment to the experimental group was associated with increased self-reported physical activity behaviors (d=0.64). On average, participants in the experimental group had an increase of 52% in their Godin Leisure-Time Exercise Questionnaire score, whereas participants in the comparison group tended to have a slight decrease in their score (Multimedia Appendix 3).

Step Count

The intraclass correlation coefficient for participants' mean daily step count per journey quartile was 0.73 (95% CI 0.64 to 0.81). The linear growth model results indicated a statistically significant group-by-time interaction for step count during the intervention (Table 3; Multimedia Appendices 4-6). The experimental group participants were more likely to increase their step counts during the intervention (d=0.28) than the comparison group participants, and this trend extended into the follow-up period (Figure 3). The linear growth model suggested that the experimental group participants tended to have a lower step count earlier in the intervention than the comparison group participants, but this difference was not statistically significant.

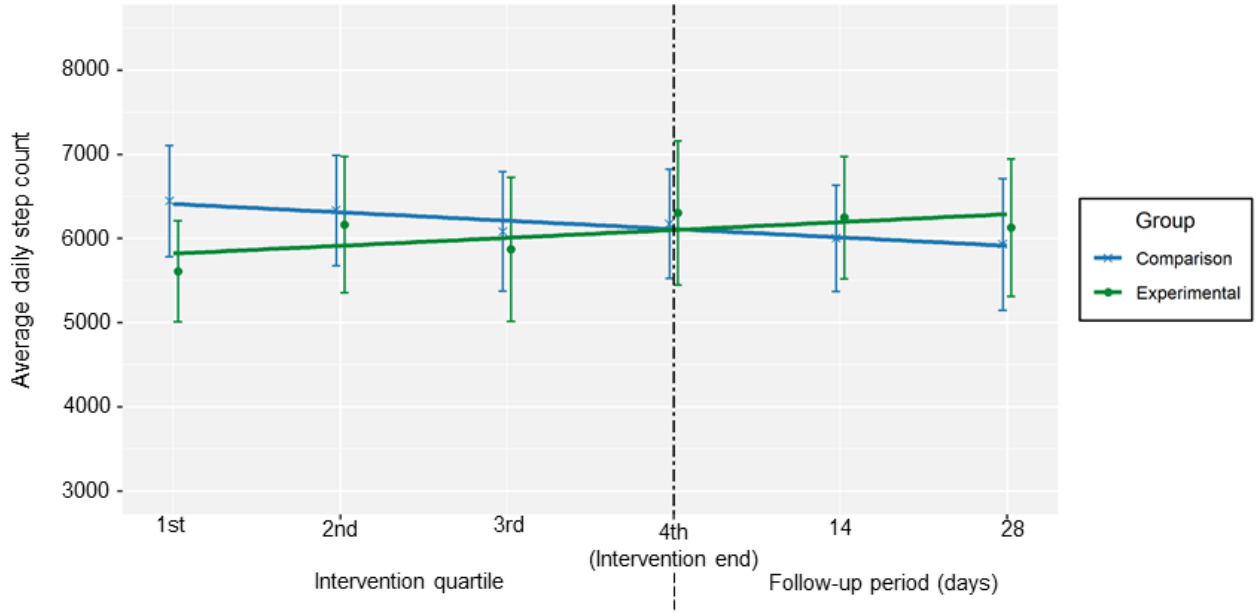
Table 3. Parameter estimates for the linear growth model of the mean daily step count per intervention quartile as a function of group assignment.

Effect	Estimate SE	SE	*P* value	95% CI
Fixed				
Intercept	6473.46	324.26	<.001	5837.91 to 7109.01
Time	–107.41	96.06	.27	–295.69 to 80.87
Group	–885.80	466.38	.06	–1799.90 to 28.30
Days of valid wear	–287.14	227.86	.21	–733.75 to 159.47
Group by time	322.08	136.77	.02	54.01 to 590.15
Random at level 2 (between persons)[a]				
Intercept	2,648,712	628,223	<.001	1,417,395 to 3,880,029
Slope for time	1284	21,090	.48	–40,052 to 42,620
Random at level 1 (within person)[b]				
Residual	1,399,430	207,992	<.001	991,766 to 1,807,094

[a]The correlation between the intercept and slope for time was 0.90.

[b]The autocorrelation was 0.21.

Figure 3. Mean device-measured daily step counts with linear trend lines by group over the study period. Quartiles are presented for the intervention period because the intervention duration differed for each pair. The median intervention period was 30 days (IQR 23-51 days).



Discussion

Principal Findings and Comparison With Prior Work

Steps2Health is a physical activity intervention that pairs game design elements with wearable mHealth technologies to target SDT constructs. In general, the results of our study support the feasibility of sending text messages for physical activity promotion contingent on real-time step counts to insufficiently active cancer survivors. One exception is that the participants' response rate to the text messages requesting a response was low; this corroborates previous literature that has identified participant engagement as a challenge to mHealth interventions [41,42]. Our results also corroborate previous studies' findings supporting the feasibility and acceptability of behavioral interventions centered on providing wearables to cancer survivors to promote their physical activity [43]. Over 96% (75/78) of the participants in this study completed the intervention, and similar to other studies, the percentage of valid wear time of the Fitbit device in this study sample was high [43,44]. Participants generally rated their satisfaction with the intervention as high but had mixed ratings for aspects of the intervention centered on targeting *relatedness*. Questionnaire data suggested that the *Steps2Health* intervention increased self-reported moderate-to-vigorous physical activity levels from pre- to postintervention. Device-measured data suggested that the *Steps2Health* intervention was associated with a greater increase in daily step count over the course of the study period;

the comparison group appeared to have a higher step count early in the intervention period, which steadily declined, whereas the experimental group had a lower step count early on, which steadily increased.

The game design elements of the *Steps2Health* intervention are in line with the recommendation by Nicholson [45] that, to foster long-term behavior change, gamified intervention approaches should be centered on providing information that engenders meaningful connections. Beyond providing performance feedback, which may not always be relevant to insufficiently active cancer survivors [29], the *Steps2Health* intervention messages were designed to build *autonomous motivation* by fostering curiosity, unexpected adventures, and playfulness in the context of cancer survivorship. The results of this study support the general acceptability of this approach and the feasibility of using the approach in real time. Survey and qualitative data from individual interviews indicated that participants found the messages targeting *autonomous motivation* to be pleasant and compelling and that they would recommend the experience to friends and family. The messages targeting *autonomy* and *competence or self-efficacy* were also well rated.

Compared with the messages targeting *autonomy* and *competence or self-efficacy*, the messages targeting *relatedness* were not as well received, and the feasibility criterion associated with the response rate to these messages was not met. Enhancing *relatedness* has the potential to promote physical activity in cancer survivors; however, the marked heterogeneity of this population introduces challenges [29,46]. Some individuals may be more receptive to having this need fulfilled by constructive social support, whereas others may benefit more from healthy competition [47]. The incorporation of more granular, individualized tailoring may enable future mHealth interventions to meet this need for a greater diversity of individuals. In this study, participants recommended providing multiple role model characters to choose from, which they felt would help to enhance relatability. Anticipating our study sample, we created a female role model, but our findings support the notion emphasized in previous literature that participant *self-tailoring* may foster individual autonomy and connection to digital interventions (eg, allowing users to choose role models whose background and story most resonate with their own) [41]. Some participants suggested including aspects of social media as a part of the intervention experience; however, this idea was met with mixed opinions and some skepticism.

One increasingly viable intervention option that may serve to enhance *relatedness* in cancer survivors in mHealth interventions to some degree is the use of more sophisticated conversational technology, such as automated conversational agents, known as chatbots. Chatbots are increasingly being used as health intervention components [48,49] and this emerging technology could help increase participant engagement and/or facilitate feelings of *relatedness* for some individuals while preserving the scalability of the intervention. *Relatedness* is theorized to influence health-related outcomes both directly and by impacting *autonomous regulation* [13]. Accordingly, targeting *relatedness* should be a priority for interventionists but the degree to which this psychological need can be met via electronic means is unclear. More research is needed to determine how mHealth interventions can meaningfully impact this construct in cancer survivors.

Research indicates that mHealth programs generally have high rates of participant discontinuation [50] and that participant engagement plays an important role in influencing the decision to continue an intervention [41,42]. In this study, Fitbit device wear was relatively high, and participants tended to rate the intervention favorably. However, the messages that requested a response were not well received. Individual interviews revealed that one potential barrier to this aspect of participant engagement was that the messages requesting a response sometimes required participants to engage in deeper reflection. This was generally perceived as inappropriate for the context in which participants tended to receive the messages (ie, when they were more active). Thus, one lesson learned from this study is to carefully consider the context in which intervention messages are to be received. We found that participants were not very likely to review past text messages, and this sentiment was echoed by participants' admission that they seldom accessed previously sent links to resources for healthy living. Future mHealth physical activity interventions could potentially increase engagement by configuring messages to be sent during lulls in physical activity that follow bouts of exercise (eg, at the conclusion of bouts of physical activity as registered by meeting a threshold level of Fitbit Active Minutes over a predetermined time period). This may be more conducive to participant engagement and serve as a useful opportunity to provide positive reinforcement and/or self-reflective, just-in-time support [51].

Our operationalization of participant response rate in this study was a summative measure of frequency. Future research should investigate what constitutes *effective engagement* or a broader conceptualization of engagement that relates more directly to desired intervention outcomes [41,42]. In this study, although the participant engagement feasibility criterion corresponding to the text message response rate was not met, participants in the experimental group tended to increase their physical activity levels more than their counterparts in the comparison group. A successful aspect of this study that helped facilitate participant engagement and high device wear was our inclusion of automated messages that were automatically sent if participants did not synchronize their Fitbit device for a prescribed period of time. The timestamp of users' last synchronization is available through the Fitbit application programming interface. In this study, we were able to use these data to help increase participants' engagement and data collection in a manner that minimized the use of study resources.

Although the experimental group reported increases in physical activity, the study results were mixed regarding whether participants felt that the *Steps2Health* messages motivated them to increase their physical activity levels. Interestingly, Zuckerman and Gal-Oz [52] noted a similar phenomenon regarding a discrepancy between the perceived and empirical effects of a gamified physical activity intervention. Participants may not necessarily attribute lifestyle changes to apparently effective gamification elements. Furthermore, we found no evidence that group assignment was associated with an increase in *autonomous regulation*. A review conducted by Johnson [53]

revealed that the effects of gamification-centered wellness interventions are strongest for physical activity–related behavioral outcomes but weaker for motivation-related cognitions. Nonetheless, both the experimental and comparison groups exhibited a small, statistically significant increase in *autonomous regulation.* The provision of a wearable device alone could impact *autonomous regulation* in this population; however, this notion should be met with caution. The provision of consumer wearable technologies alone may undermine physical activity–related motivations in some populations [54] by emphasizing less autonomous forms of motivation, inducing anxiety and stigma, or by imposing unattainable or inappropriate norms [55].

Our findings indicated that the *Steps2Health* intervention was associated with an increase in self-reported exercise and device-measured step count during the intervention. Survey analyses provided evidence for a medium effect size from baseline to follow-up that is consistent with other mHealth-based behavioral interventions for promoting cancer survivors' physical activity [56,57]. This finding was supplemented by an analysis of device-measured step counts over the intervention period. The comparison group participants' step count was initially high but trended steadily downward. This is concordant with a waning *novelty effect* associated with physical activity device technologies [12]. The gamification aspects of the *Steps2Health* intervention may have influenced participants to gradually increase their step count. Future studies should further investigate these dynamics, as long-term physical activity maintenance is needed to best realize the diverse health benefits of physical activity. In some cases, it may be prudent to delay game design intervention components to capitalize on their effects and the *novelty effect* associated with the provision of physical activity tracking technologies. In fact, the Maintain IT Model of health behavior change posits that interventions centered on self-regulation may be appropriate for health behavior initiation, whereas interventions centered on precipitating a more centered and empowered state, consistent with SDT, may be more effective for leading to long-term health behavior maintenance [58].

Study Limitations

The establishment of the acceptability and feasibility of digital behavior change interventions is critically important, given their high initial costs and high rate of discontinuation of use. Furthermore, the CONSORT (Consolidated Standards of Reporting Trials) 2010 Statement Extension to Randomized Pilot and Feasibility Trials outlines the need for prespecified assessments of pilot trial objectives [59]. Unfortunately, there is a lack of consensus on the established criteria to aid researchers in determining the feasibility and acceptability of digital behavior change interventions. In this study, we set an a priori criteria for assessing the feasibility of *Steps2Health* based on the intervention's unique components (eg, text messaging contingent on physical activity tracker device–derived data), data observed in comparable studies (eg, for physical activity tracker wear), and the theoretical orientation of our intervention (ie, the emphasis on intrinsic motivation in SDT). However, we acknowledge that these criteria, and the thresholds we set for them, may not be applicable to other studies (or necessarily the ideal benchmarks for evaluating the feasibility of *Steps2Health*). Establishing universal quantitative criteria and corresponding thresholds may be a challenge, given the heterogeneous and rapidly evolving nature of digital behavior change interventions, but establishing basic metrics for commonly occurring physical activity intervention components may facilitate the conduct of rigorous research. Future research on this topic could make a meaningful contribution to the literature.

This study was limited by its small sample size and the use of convenience sampling recruitment methods. Participants may have been particularly motivated to change their physical activity levels from the outset, which may have interacted with the intervention components. The generalizability of the study's findings is limited by the fact that the study sample was relatively well educated and mostly female. Another limitation of this study is that we did not audio record the in-depth interviews and used field notes instead. Loss of details from field notes may have increased the risk of bias in the qualitative analysis. Furthermore, although every effort was made to capture insightful quotations accurately, some errors may have been introduced in this process. Nevertheless, associated threats to validity were reduced by our relatively straightforward applied research questions, short interviews, study staff members exclusively dedicated to taking field notes, and a research protocol specifying that the interviewer separately provided field notes immediately after all interviews. The measurement of physical activity via self-report has known limitations, and this is also true for consumer-grade wearable devices. Fitbit devices tend to overestimate step counts in free-living conditions, particularly for individuals with chronic disease and/or mobility limitations [38]. However, this study's limitations are offset by its notable strengths, including a randomized controlled design, a novel intervention approach, minimal loss to follow-up, the use of both quantitative and qualitative methods, and intervention qualities that lend themselves readily to widespread dissemination.

Conclusions

The findings of this study support the feasibility and acceptability of using the gamification of real-time step counts to increase cancer survivors' physical activity. Both the experimental and comparison groups increased their *autonomous regulation* for physical activity, and assignment to the experimental group was associated with increased physical activity. Coupling game design elements with wearable technologies is technically feasible and acceptable to cancer survivors and is potentially effective. More research is needed to develop these approaches as they have the potential to have a meaningful impact on public health.

Acknowledgments

This research was made possible by the Center for Energy Balance in Cancer Prevention and Survivorship and by the Duncan Family Institute for Cancer Prevention and Risk Assessment, both at the University of Texas MD Anderson Cancer Center. MR was supported by the National Cancer Institute of the National Institutes of Health (F31 CA236433). This study used the assessment, intervention, and measurement (AIM) core by MD Anderson, a shared resource supported by the National Institutes of Health through MD Anderson's Cancer Center Support Grant (P30 CA016672). This study was supported by the Center for Energy Balance in Cancer Prevention and Survivorship, the Duncan Family Institute for Cancer Prevention and Risk Assessment, and a donation from Jason's Deli. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. EL was supported by a Mentored Research Scholar Grant in Applied and Clinical Research from the American Cancer Society (MRSG-14-165-01-CPPB). YL was supported by a faculty fellowship from the Duncan Family Institute for Cancer Prevention and Risk Assessment.

Authors' Contributions

MR contributed to the conception of the study design, data collection, quantitative and qualitative data analysis, and drafting of the paper and provided approval of the version to be published. EL contributed to the conception of the study design, provided critical revision of the paper, and provided approval of the version to be published. YL contributed to quantitative data analysis, provided critical revision of the paper, and provided approval of the version to be published. MB contributed to the conception of the study design, data collection, and qualitative data analysis and provided approval of the version to be published. KBE contributed to the conception of the study design, data collection, quantitative and qualitative data analysis; provided critical revision of the paper; and provided approval of the version to be published.

Conflicts of Interest

None declared.

This randomized study was not registered. The authors explained that their study was "not an applicable clinical trial and did not require registration in the Clinical Trials Registry." The editor granted an exception of ICMJE rules for prospective registration of randomized trials because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness.

Multimedia Appendix 1
Steps2Health interview guide.
[DOCX File , 20 KB - mhealth_v8i11e18364_app1.docx]

Multimedia Appendix 2
Conceptual model for Steps2Health intervention.
[PNG File , 40 KB - mhealth_v8i11e18364_app2.png]

Multimedia Appendix 3
Mean Godin Leisure-Time Physical Activity Questionnaire scores by group at baseline and follow-up. Error bars indicate 95% CIs.
[PNG File , 23 KB - mhealth_v8i11e18364_app3.png]

Multimedia Appendix 4
Spaghetti plots of mean (thick) and subject-specific (thin) step count trajectories for the comparison and experimental groups during the intervention period.
[PNG File , 39 KB - mhealth_v8i11e18364_app4.png]

Multimedia Appendix 5
Raw and fitted step count trajectories for individual participants in the experimental group.
[PNG File , 84 KB - mhealth_v8i11e18364_app5.png]

Multimedia Appendix 6
Raw and fitted step count trajectories for individual participants in the comparison group.
[PNG File , 85 KB - mhealth_v8i11e18364_app6.png]

References

1. Ballard-Barbash R, Friedenreich CM, Courneya KS, Siddiqi SM, McTiernan A, Alfano CM. Physical activity, biomarkers, and disease outcomes in cancer survivors: a systematic review. J Natl Cancer Inst 2012 Jun 6;104(11):815-840 [FREE Full text] [doi: 10.1093/jnci/djs207] [Medline: 22570317]
2. Mishra SI, Scherer RW, Geigle PM, Berlanstein DR, Topaloglu O, Gotay CC, et al. Exercise interventions on health-related quality of life for cancer survivors. Cochrane Database Syst Rev 2012 Aug 15(8):CD007566 [FREE Full text] [doi: 10.1002/14651858.CD007566.pub2] [Medline: 22895961]
3. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. J Am Med Assoc 2018 Nov 20;320(19):2020-2028. [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
4. Tarasenko Y, Chen C, Schoenberg N. Self-reported physical activity levels of older cancer survivors: results from the 2014 national health interview survey. J Am Geriatr Soc 2017 Feb;65(2):e39-e44. [doi: 10.1111/jgs.14589] [Medline: 27943255]
5. Loprinzi PD, Lee H, Cardinal BJ. Objectively measured physical activity among US cancer survivors: considerations by weight status. J Cancer Surviv 2013 Sep;7(3):493-499. [doi: 10.1007/s11764-013-0293-7] [Medline: 23686709]
6. Thraen-Borowski KM, Gennuso KP, Cadmus-Bertram L. Accelerometer-derived physical activity and sedentary time by cancer type in the United States. PLoS One 2017;12(8):e0182554 [FREE Full text] [doi: 10.1371/journal.pone.0182554] [Medline: 28806753]
7. Milne HM, Guilfoyle A, Gordon S, Wallman KE, Courneya KS. Personal accounts of exercise and quality of life from the perspective of breast cancer survivors. Qual Life Res 2007 Nov;16(9):1473-1481. [doi: 10.1007/s11136-007-9251-z] [Medline: 17717726]
8. Smith L, Croker H, Fisher A, Williams K, Wardle J, Beeken RJ. Cancer survivors' attitudes towards and knowledge of physical activity, sources of information, and barriers and facilitators of engagement: A qualitative study. Eur J Cancer Care (Engl) 2017 Jul;26(4):-. [doi: 10.1111/ecc.12641] [Medline: 28135016]
9. Haberlin C, O'Dwyer T, Mockler D, Moran J, O'Donnell DM, Broderick J. The use of eHealth to promote physical activity in cancer survivors: a systematic review. Support Care Cancer 2018 Oct;26(10):3323-3336. [doi: 10.1007/s00520-018-4305-z] [Medline: 29909476]
10. Schaffer K, Panneerselvam N, Loh KP, Herrmann R, Kleckner IR, Dunne RF, et al. Systematic review of randomized controlled trials of exercise interventions using digital activity trackers in patients with cancer. J Natl Compr Canc Netw 2019 Jan;17(1):57-63 [FREE Full text] [doi: 10.6004/jnccn.2018.7082] [Medline: 30659130]
11. Thomas R. 25% of Users Abandon Apps After One Use. Localytics. 2019. URL: https://tinyurl.com/y6qbn98l [accessed 2020-10-16]
12. Shin G, Feng Y, Jarrahi M, Gafinowitz N. Beyond novelty effect: a mixed-methods exploration into the motivation for long-term activity tracker use. JAMIA Open 2019 Apr;2(1):62-72 [FREE Full text] [doi: 10.1093/jamiaopen/ooy048] [Medline: 31984346]
13. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
14. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012 Jun 22;9:78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
15. Milne HM, Wallman KE, Guilfoyle A, Gordon S, Corneya KS. Self-determination theory and physical activity among breast cancer survivors. J Sport Exerc Psychol 2008 Feb;30(1):23-38. [doi: 10.1123/jsep.30.1.23] [Medline: 18369241]
16. Mayer DK, Landucci G, Awoyinka L, Atwood AK, Carmack CL, Demark-Wahnefried W, et al. SurvivorCHESS to increase physical activity in colon cancer survivors: can we get them moving? J Cancer Surviv 2018 Feb;12(1):82-94 [FREE Full text] [doi: 10.1007/s11764-017-0647-7] [Medline: 28994035]
17. Mendoza JA, Baker KS, Moreno MA, Whitlock K, Abbey-Lambertz M, Waite A, et al. A Fitbit and Facebook mHealth intervention for promoting physical activity among adolescent and young adult childhood cancer survivors: A pilot study. Pediatr Blood Cancer 2017 Dec;64(12):-. [doi: 10.1002/pbc.26660] [Medline: 28618158]
18. Przybylski A, Rigby C, Ryan R. A motivational model of video game engagement. Rev General Psychol 2010 Jun;14(2):154-166. [doi: 10.1037/a0019440]
19. Lyons EJ, Tate DF, Ward DS, Ribisl KM, Bowling JM, Kalyanaraman S. Engagement, enjoyment, and energy expenditure during active video game play. Health Psychol 2014 Feb;33(2):174-181 [FREE Full text] [doi: 10.1037/a0031947] [Medline: 23527520]
20. Sailer M, Hense J, Mayr S, Mandl H. How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction. Comput Hum Behav 2017 Apr;69:371-380. [doi: 10.1016/j.chb.2016.12.033]
21. Deterding S. The lens of intrinsic skill atoms: a method for gameful design. Hum–Comput Interac 2015 May 15;30(3-4):294-335. [doi: 10.1080/07370024.2014.993471]
22. 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;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]

23. Lucero A, Holopainen J, Ollila E, Suomela R, Karapanos E. The playful experiences (PLEX) framework as a guide for expert evaluation. In: The 6th International Conference on Designing Pleasurable Products and Interfaces. 2013 Presented at: PLEX'13; September, 2013; Newcastle upon Tyne. [doi: 10.1145/2513506.2513530]
24. Baranowski T, Lyons EJ, Thompson D. Experimental design to systematically develop a knowledge base for effective games for health. Games Health J 2019 Oct;8(5):307-312. [doi: 10.1089/g4h.2019.0057] [Medline: 31157990]
25. Hayes SC. Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies - republished article. Behav Ther 2016 Nov;47(6):869-885. [doi: 10.1016/j.beth.2016.11.006] [Medline: 27993338]
26. Miller W, Rollnick S. Motivational interviewing Helping People Change. New York, USA: The Guilford Press; 2013.
27. Schmitz KH, Courneya KS, Matthews C, Demark-Wahnefried W, Galvão DA, Pinto BM, American College of Sports Medicine. American College of Sports Medicine roundtable on exercise guidelines for cancer survivors. Med Sci Sports Exerc 2010 Jul;42(7):1409-1426. [doi: 10.1249/MSS.0b013e3181e0c112] [Medline: 20559064]
28. Shephard R. Godin Leisure-Time Exercise Questionnaire. Medicine&amp Science in Sports &amp Exercise 1997 Jun;29(Supplement):36-38 [FREE Full text] [doi: 10.1097/00005768-199706001-00009]
29. Robertson MC, Tsai E, Lyons EJ, Srinivasan S, Swartz MC, Baum ML, et al. Mobile health physical activity intervention preferences in cancer survivors: a qualitative study. JMIR Mhealth Uhealth 2017 Jan 24;5(1):e3 [FREE Full text] [doi: 10.2196/mhealth.6970] [Medline: 28119278]
30. Brickwood K, Watson G, O'Brien J, Williams AD. Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Apr 12;7(4):e11819 [FREE Full text] [doi: 10.2196/11819] [Medline: 30977740]
31. Peddle CJ, Plotnikoff RC, Wild TC, Au H, Courneya KS. Medical, demographic, and psychosocial correlates of exercise in colorectal cancer survivors: an application of self-determination theory. Support Care Cancer 2008 Jan;16(1):9-17. [doi: 10.1007/s00520-007-0272-5] [Medline: 17569994]
32. Schmitz KH, Campbell AM, Stuiver MM, Pinto BM, Schwartz AL, Morris GS, et al. Exercise is medicine in oncology: engaging clinicians to help patients move through cancer. CA Cancer J Clin 2019 Nov;69(6):468-484 [FREE Full text] [doi: 10.3322/caac.21579] [Medline: 31617590]
33. Liao Y, Robertson MC, Winne A, Wu IH, Le TA, Balachandran DD, et al. Investigating the within-person relationships between activity levels and sleep duration using Fitbit data. Transl Behav Med 2020 Jul 15 epub ahead of print. [doi: 10.1093/tbm/ibaa071] [Medline: 32667039]
34. Chu AH, Ng SH, Paknezhad M, Gauterin A, Koh D, Brown MS, et al. Comparison of wrist-worn Fitbit Flex and waist-worn ActiGraph for measuring steps in free-living adults. PLoS One 2017;12(2):e0172535 [FREE Full text] [doi: 10.1371/journal.pone.0172535] [Medline: 28234953]
35. Tudor-Locke C, Barreira TV, Schuna JM. Comparison of step outputs for waist and wrist accelerometer attachment sites. Med Sci Sports Exerc 2015 Apr;47(4):839-842. [doi: 10.1249/MSS.0000000000000476] [Medline: 25121517]
36. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364 [FREE Full text] [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]
37. Collins JE, Yang HY, Trentadue TP, Gong Y, Losina E. Validation of the Fitbit Charge 2 compared to the ActiGraph GT3X+ in older adults with knee osteoarthritis in free-living conditions. PLoS One 2019;14(1):e0211231 [FREE Full text] [doi: 10.1371/journal.pone.0211231] [Medline: 30699159]
38. 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]
39. Silva SS, Jayawardana MW, Meyer D. Statistical methods to model and evaluate physical activity programs, using step counts: a systematic review. PLoS One 2018;13(11):e0206763 [FREE Full text] [doi: 10.1371/journal.pone.0206763] [Medline: 30388164]
40. Bolger N, Laurenceau J. In: Kenny D, Little T, editors. Intensive Longitudinal Methods: an Introduction to Diary and Experience Sampling Research. New York, USA: Guilford Press; 2013.
41. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
42. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
43. Hartman SJ, Marinac CR, Bellettiere J, Godbole S, Natarajan L, Patterson RE, et al. Objectively measured sedentary behavior and quality of life among survivors of early stage breast cancer. Support Care Cancer 2017 Aug;25(8):2495-2503 [FREE Full text] [doi: 10.1007/s00520-017-3657-0] [Medline: 28326507]
44. Maxwell-Smith C, Hince D, Cohen PA, Bulsara MK, Boyle T, Platell C, et al. A randomized controlled trial of WATAAP to promote physical activity in colorectal and endometrial cancer survivors. Psychooncology 2019 Jul;28(7):1420-1429. [doi: 10.1002/pon.5090] [Medline: 30980691]

45. Nicholoson S. A RECIPE for Meaningful Gamification. In: Reiners T, Wood LC, editors. Gamification in Education and Business. Cham: Springer; 2015:1-20.
46. Tsai E, Robertson MC, Lyons EJ, Swartz MC, Basen-Engquist K. Physical activity and exercise self-regulation in cancer survivors: a qualitative study. Psychooncology 2018 Feb;27(2):563-568 [FREE Full text] [doi: 10.1002/pon.4519] [Medline: 28763133]
47. Zhang J, Brackbill D, Yang S, Becker J, Herbert N, Centola D. Support or competition? How online social networks increase physical activity: a randomized controlled trial. Prev Med Rep 2016 Dec;4:453-458 [FREE Full text] [doi: 10.1016/j.pmedr.2016.08.008] [Medline: 27617191]
48. Pereira J, Díaz O. Using health chatbots for behavior change: a mapping study. J Med Syst 2019 Apr 4;43(5):135. [doi: 10.1007/s10916-019-1237-1] [Medline: 30949846]
49. Kowatsch T, Nißen M, Rüegger D, Stieger M, Flückiger C, Allemand M, et al. The Impact of Interpersonal Closeness Cues in Text-based Healthcare Chatbots on Attachment Bond and the Desire to Continue Interacting: an Experimental Design. In: 26th European Conference on Information Systems. 2018 May Presented at: CIS'18; June 23, 2018; Portsmouth, UK URL: https://www.alexandria.unisg.ch/publications/254284
50. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
51. Nahum-Shani I, Smith S, Spring B, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
52. Zuckerman O, Gal-Oz A. Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. Pers Ubiquit Comput 2014 Jul 5;18(7):1705-1719 [FREE Full text] [doi: 10.1007/s00779-014-0783-2]
53. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
54. Kerner C, Goodyear V. The motivational impact of wearable healthy lifestyle technologies: a self-determination perspective on fitbits with adolescents. Am J Health Educ 2017 Jul 28;48(5):287-297 [FREE Full text] [doi: 10.1080/19325037.2017.1343161]
55. Owens J, Cribb A. 'My fitbit thinks I can do better!' do health promoting wearable technologies support personal autonomy? PhilosTechnol 2017 Jun 6;32(1):23-38 [FREE Full text] [doi: 10.1007/s13347-017-0266-2]
56. Kiss N, Baguley BJ, Ball K, Daly RM, Fraser SF, Granger CL, et al. Technology-supported self-guided nutrition and physical activity interventions for adults with cancer: systematic review. JMIR Mhealth Uhealth 2019 Feb 12;7(2):e12281 [FREE Full text] [doi: 10.2196/12281] [Medline: 30747720]
57. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017 Dec;11(6):704-719 [FREE Full text] [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]
58. Caldwell AE, Masters KS, Peters JC, Bryan AD, Grigsby J, Hooker SA, et al. Harnessing centred identity transformation to reduce executive function burden for maintenance of health behaviour change: the Maintain IT model. Health Psychol Rev 2018 Sep;12(3):231-253 [FREE Full text] [doi: 10.1080/17437199.2018.1437551] [Medline: 29402182]
59. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Br Med J 2016 Oct 24;355:i5239. [doi: 10.1136/bmj.i5239] [Medline: 27777223]

Abbreviations

BREQ-3: Behavioral Regulation for Exercise Questionnaire-3
mHealth: mobile health
MMS: multimedia messaging service
SDT: Self-Determination Theory

Edited by G Eysenbach; submitted 21.02.20; peer-reviewed by N Ridgers, R Rahman, C Romanzini, L Grepo, AS Lu; comments to author 29.06.20; revised version received 24.08.20; accepted 15.09.20; published 24.11.20.

Please cite as:
Robertson MC, Lyons EJ, Liao Y, Baum ML, Basen-Engquist KM
Gamified Text Messaging Contingent on Device-Measured Steps: Randomized Feasibility Study of a Physical Activity Intervention for Cancer Survivors
JMIR Mhealth Uhealth 2020;8(11):e18364
URL: https://mhealth.jmir.org/2020/11/e18364
doi:10.2196/18364
PMID:33231551

Original Paper

Mobile Phone Apps in Australia for Improving Pregnancy Outcomes: Systematic Search on App Stores

Loretta M Musgrave[1,2], MEd; Nathalie V Kizirian[2], PhD; Caroline S E Homer[3,4], PhD; Adrienne Gordon[2,5], PhD

[1]Centre for Midwifery, Child and Family Health, University of Technology Sydney, Ultimo NSW, Australia

[2]Faculty of Medicine and Health, Charles Perkins Centre, University of Sydney, Camperdown NSW, Australia

[3]Burnet Institute, Melbourne VIC, Australia

[4]Centre for Midwifery, Child, and Family Health, University of Technology Sydney, Ultimo NSW, Australia

[5]Sydney Local Health District, NSW Health, Camperdown NSW, Australia

Corresponding Author:
Loretta M Musgrave, MEd
Centre for Midwifery, Child and Family Health
University of Technology Sydney
Building 10, Level 11
235 Jones St
Ultimo NSW
Australia
Phone: 61 421633406
Email: loretta.musgrave@uts.edu.au

Abstract

Background: Women are increasingly turning to mobile health platforms to receive health information and support in pregnancy, yet the content of these platforms vary. Although there is great potential to influence health behaviors, little research has assessed the quality of these platforms or their ability to change behavior. In recent years, validated tools to assess app quality have become available.

Objective: To identify and assess the quality and ongoing popularity of the top 10 freely available pregnancy apps in Australia using validated tools.

Methods: A systematic search on app stores to identify apps was performed. A Google Play search used subject terms pregnancy, parenting, and childbirth; the iTunes search used alternative categories medical and health and fitness. The top 250 apps from each store were cross-referenced, and the top 100 found in both Google Play and iTunes were screened for eligibility. Apps that provided health information or advice for pregnancy were included. Excluded apps focused on nonhealth information (eg, baby names). The top 10 pregnancy apps were assessed using the Mobile App Rating Scale (MARS). A comparative analysis was conducted at 2 time points over 2 years to assess the ongoing popularity of the apps. The MARS score was compared to the download and star rating data collected from iTunes and Google Play in 2017 and 2019. Health behaviors including breastfeeding, healthy pregnancy weight, and maternal awareness of fetal movements were reviewed for apparent impact on the user's knowledge, attitudes, and behavior change intentions using the MARS perceived impact section and the Coventry, Aberdeen, and London—Refined (CALO-RE) taxonomy.

Results: A total of 2052 free apps were screened for eligibility, 1397 were excluded, and 655 were reviewed and scored. The top 10 apps were selected using download numbers and star ratings. All 10 apps were suboptimal in quality, practicality, and functionality. It was not possible to identify a primary purpose for all apps, and there was overlap in purpose for many. The mean overall MARS app quality score across all 10 apps was 3.01 (range 1.97-4.40) in 2017 and 3.40 (range 2.27-4.44) in 2019. A minority of apps scored well for perceived impact on health behavior using the MARS tool. Using the CALO-RE 40 item taxonomy, the number of behavior change techniques used was low. The mean number of behavior change techniques for breastfeeding was 5 (range 2-11), for pregnancy weight was 4 (range 2-12), and for maternal awareness of fetal movements was 5 (range 2-8).

Conclusions: This review provides valuable information to clinicians and consumers about the quality of apps currently available for pregnancy in Australia. Consideration is needed regarding the regulation of information and the potential opportunity to incorporate behavior change techniques to improve maternal and fetal outcomes.

(JMIR Mhealth Uhealth 2020;8(11):e22340) doi:10.2196/22340

KEYWORDS

smartphone apps; mobile phone; pregnancy; health behavior change; MARS tool; CALO-RE taxonomy; pregnancy outcomes; quality assessment methods

Introduction

Smartphone ownership and app use in Australia are high, with 81% of people possessing a smartphone, and 97% of mobile consumers aged between 18 and 34 years [1]. In 2019, mobile phones were the most common device used to access the internet (87%), followed by a laptop (69%), then tablets (56%) [2]. The most recent Australian data suggest that 46% of internet users access the internet for health services; this is an increase from 22% in 2014-2015 [3]. It has been estimated that up to 1 in 4 Australians use their smartphones to access health-related apps to support healthy behaviors [4].

Pregnant women are increasingly turning to mobile health (mHealth) to receive health information and support rather than relying on face-to-face and paper-based delivery methods [5-11]. This use of mobile health apps during pregnancy provides a unique window of opportunity—a teachable moment—when women are often more motivated to optimize health and change their lifestyle [12,13]. Apps also have the potential to act as a platform for specific pregnancy behavior change interventions, such as maternal awareness of decreased fetal movements, maternal weight monitoring, and breastfeeding [14-17]. A recent systematic review [18] found limited data of the effects of mobile app interventions during pregnancy on maternal knowledge and behavior change. This review [18] concluded that well-designed studies are needed to evaluate apps. App developers should include women in co-design, implementation, and evaluation phases of development. This was further supported by a systematic review [19] that aimed to evaluate usability (feasibility and acceptability) as well as the effectiveness of lifestyle and medical apps in supporting health care during pregnancy in high-income countries. The review [19] concluded that further evidence is needed before such apps are implemented in health care. For apps to be used as an adjunct to health care, issues related to the accuracy of the information, privacy, and security also need to be addressed [19].

Health behaviors such as maternal awareness of decreased fetal movements, maintaining a healthy weight in pregnancy, and breastfeeding are modifiable behaviors with known benefits for both mothers and babies. Globally, apps have been used to address such behaviors; however, further evidence is needed to establish if these apps have an impact on pregnancy outcomes [18]. In Australia, such apps have been assessed in a research setting, including Growing Healthy, which provides information on healthy infant feeding [20], and the My Baby's Movements app, which provides information about normal fetal movements and has a tracking tool [21]. In 2018, a quasi-experimental study [20] was conducted to describe the effects of Growing Healthy on parental feeding practices, infant food preferences, and infant satiety responsiveness; the authors concluded that mHealth design and delivery characteristics that impact on infant feeding practices need further research. My Baby's Movements has also been tested in a randomized controlled trial [21]; results are not yet published.

Women who are overweight or obese have an increased risk of pregnancy complications [22]. Pregnancy weight gain can be addressed through lifestyle and dietary interventions [23]. Institute of Medicine weight gain recommendations [24], National Institute for Health and Clinical Excellence guidelines [25], and Australian dietary and physical activity guidelines are referenced and recommended in national Clinical Practice Guidelines for pregnancy care [26]. A recent systematic review of nutritional information available to pregnant women on smartphones in the United Kingdom found that apps do not consistently provide useful or accurate nutritional information [27].

The health benefits of exclusive breastfeeding for 6 months for mothers and infants are well established [28]. A 2019 systematic review [29] of digital interventions that support breastfeeding found that there is potential to improve breastfeeding outcomes. A recent cohort study [30] conducted in the United Kingdom evaluated the effectiveness of the Baby Buddy app; Baby Buddy is an mHealth intervention that is available on the UK National Health Service Library that supports and guides women through pregnancy and the first 6 months of their child's life. The posthoc analysis of this study suggested that Baby Buddy app users were more likely to report exclusively breastfeeding or ever breastfeeding [30].

This study aimed to identify and review the top 10 pregnancy apps available in Australia over 2 years using validated tools to assess the quality and perceived impact of 3 important pregnancy health behaviors.

Methods

Study Design

This review used a stepwise systematic approach to identify, select, assess, and evaluate the 10 most popular pregnancy apps in Australia from November 2017 to October 2019. We assessed their quality and use of behavior change techniques for 3 specified behaviors— maternal awareness of decreased fetal movement, managing weight in pregnancy, and breastfeeding—using validated tools [31,32].

Step 1: Selection of Smartphone Apps

Apps were identified using a search strategy developed using PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocol) guidelines [33] for reporting systematic reviews evaluating health care interventions. Both iTunes (Apple Inc, Australia) and Google Play (Google Inc, Australia) were searched using a set of terms developed for each online app store. The searches were conducted on the authors' smartphones. Google Play search terms included *pregnancy*, *parenting*, and *childbirth*. Categories searched in iTunes were

medical and *health and fitness*. The top 250 apps from each store that met the search terms were cross-referenced to find the top 100 available in both Google Play and iTunes. These 100 apps were then screened for eligibility. Apps were screened for relevance based on the inclusion criteria, using the information provided in the app store description. The top 10 apps were then selected using the download numbers and star ratings provided in each store. A comparative analysis of the top 10 apps was conducted at 2 time points, in November 2017 and in October 2019. Discrepancies regarding the selected apps were discussed and resolved by the review team.

Apps were included in the search that satisfied the following criteria: available free (with or without in-app purchase) AND modifiable or interactive AND provided general pregnancy information or education. In addition, there was a criterion that the app either aimed to support targeted behavior change in pregnancy such as healthy diet and exercise or aimed to support general well-being and disease prevention in pregnancy including mental health. Apps were excluded in the search if they satisfied any of the following criteria: a cost was involved, the app was not available in both Australia iTunes or Google Play stores, the app was not available in English, the app was not designed for interactive use, the app was primarily designed to track or assist with contraception or fertility, the app was classified as a game or entertainment only and was not designed for education or information delivery (eg, baby names), or the app was designed primarily for use by other consumers (such as health care professionals, women's partners).

Step 2: Evaluation of Smartphone Apps

App Classification and Quality

Two reviewers classified and evaluated the quality of the top 10 pregnancy apps using the Mobile Application Rating Scale (MARS) [31]. The MARS tool was chosen as it has proven reliability through test-retest studies and has excellent internal consistency [31]. MARS has been validated for health applications and has been used in several studies, for example, pregnancy-specific nutrition apps [34,35], medication adherence [36], apps for treatment of speech disorders in children [37] and pain management [38]. Using the descriptive information provided by each app, the reviewers identified the focus and theoretical background (or strategies) used by the app developers. Affiliations, technical aspects, and target age groups were also examined. The 23-item tool has 4 objective quality subscales and 1 subjective quality rating scale. A 5-point rating scale (1, inadequate, to 5, excellent) was used for each of the 4 objective subscales: engagement, functionality, aesthetics, and information quality. A mean score was calculated that ranged from 1-5. A score of 5 denoted excellent quality, and a score of 1 indicated poor app quality [31]. The overall app quality score was calculated using the scores for the 4 domains. MARS has been designed in this way so that the total score can be directly translated to a star rating, and therefore, can be easily compared with app stores. Each app was assessed for subjective quality using a 5-point scale and calculated mean. The 4 subjective items were potential benefit, use, cost, and overall personal star rating (a score of 5 denoted "One of the best apps I've used" and a score of 1 denoted "One of the worst apps I've used") [31]. Reviewers used each app for at least 10 minutes and assessed how easy the app was to use. The 10 included apps were assessed on both iOS and Android devices to determine if there was any variance in functionality or usability between the different platforms. Data related to app settings, developer information (affiliations), external links, and security features were also reviewed.

Comparative Analysis

At 2 time points, 2 years apart, a comparative analysis of the MARS scores, downloads and star ratings of the top 10 apps was conducted to assess the ongoing quality and popularity of the selected apps. Data were collected from iTunes and Google Play on November 3, 2017 and October 5, 2019. This was done to assess whether the quality or download rating had changed.

Step 3: Analysis of Behavior Change Techniques Used

The 3 prespecified target health behaviors were found in all 10 apps. These behaviors were assessed in 2019 using both the additional component of the MARS tool [31], perceived impact, and the Coventry, Aberdeen, and London—Refined (CALO-RE) taxonomy [32]. This was undertaken to compare which apps could be effective in modifying behavior change. Content related to the behaviors was reviewed in each app to assess potential impact on user awareness, knowledge, attitudes, intentions to change, help-seeking, and behavior change and was documented using a 5-point rating scale from 1 (strongly disagree) to 5 (strongly agree) [31]. To assess the perceived impact, we used a method described by Furlong et al [37]. We identified best-practice principles for decreased fetal movement awareness, weight management, and breastfeeding as well as the intervention techniques used by the apps (eg, self-monitoring, instruction on how to perform the behavior, and information on consequences of behavior). A mean score was then calculated [37]. The CALO-RE tool was chosen as an adjunct to MARS perceived impact as it has been used successfully for reporting, evaluating, and implementing physical activity, healthy eating, and lifestyle interventions [34]. CALO-RE is a systematic way to apply evidence and theory linked to behavior change using a taxonomy consisting of 40 behavior change techniques items [32]. Each app was reviewed for all 3 target behaviors. Definitions were used to accurately describe the behavior change techniques such as goal setting, action planning, and barrier identification [32]. Using a method described by Brown et al [34], we assessed the frequency of behavior change technique inclusion in all 3 behaviors in all 10 apps. To do this, we reviewed app content, assigned individual behavior change techniques as defined by Michie et al [32] and calculated a score out of 40 possible behavior change techniques [34]. We repeated this process for each behavior.

Fetal Movement Awareness

Information about fetal well-being and advice given regarding decreased fetal movement were examined using the Australian Safer Baby Bundle Handbook and Resource Guide [39]. Using MARS and CALO-RE, we considered the potential impact that each app would have on changing women behaviors toward monitoring fetal movements and the likelihood that they would act on concerns and contact a health care provider.

Healthy Weight in Pregnancy

Advice on gestational weight gain included in apps was reviewed against US Institute of Medicine weight gain recommendations, the UK National Institute for Health and Clinical Excellence weight management guidelines for pregnancy, and the Australian clinical practice guidelines for pregnancy care.

Breastfeeding

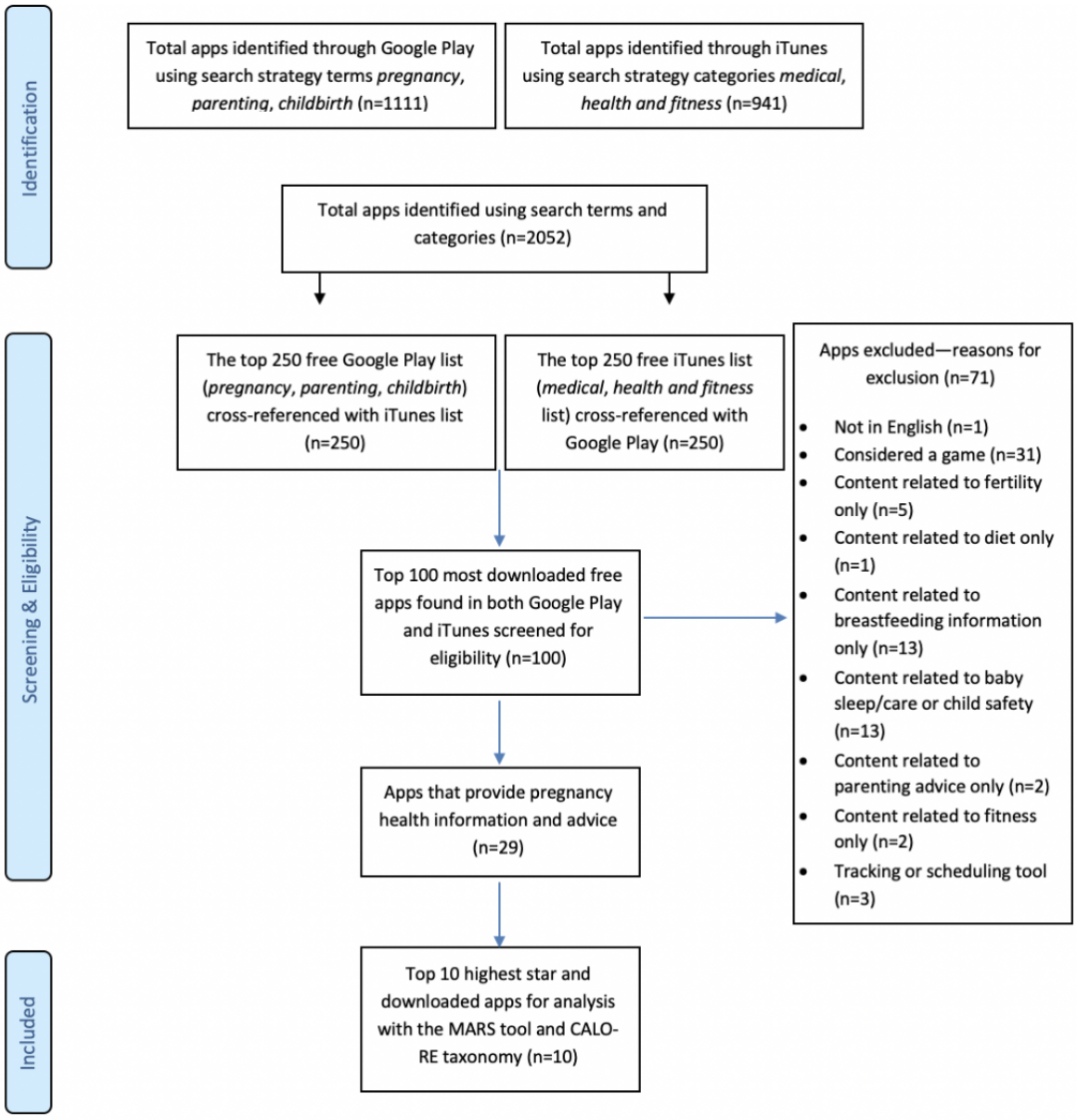
Breastfeeding content was compared for alignment with the World Health Organization (WHO) Breastfeeding Friendly Hospital Initiative 10 Steps to Successful Breastfeeding [40] and International Code of Marketing of Breast-Milk Substitutes [41]. Information, images, advertising, and sponsorship of each app were reviewed and evidence of breaches of the WHO Code was collated alongside the MARS and CALO-RE data.

Results

Step 1: Selection of Smartphone Apps

In November 2017, a total of 2052 apps were identified. Of these, 1111 apps were found only in Google Play, and 941 were found only in iTunes. Due to the volume of apps, the top 250 free apps in both stores were cross-referenced to identify the top 100 most downloaded. A total of 71 apps were excluded, and 29 met inclusion criteria. Of these 29 apps, the 10 apps with the highest number of downloads and star ratings in both stores were identified (Figure 1). The following apps were assessed for quality: Ovia Pregnancy Tracker, I'm Expecting, Baby Centre, Pregnancy +, Glow, What to Expect, Baby Bump Pregnancy Pro, Sprout Pregnancy, Week by Week, The Bump. In 2019, the top 10 apps identified from 2017 were again searched for in both stores to conduct a comparative analysis of download and star ratings at 2 time points.

Figure 1. Flowchart of the app selection process.



Step 2: Evaluation of Selected Smartphone Apps

App Classification

The focus, theoretical background, and strategies used in each app were difficult to ascertain. Affiliations and sources of funding information available indicated that all 10 were commercially developed. All apps lacked transparency regarding the details of funding. The review team was unable to determine the target age groups as these were not specified. The app description, design, functionality, and content were designed to appeal to women of reproductive age; however, some had user pathways for partners and carers. Of the 10 apps, 7 had reminders, 7 allowed password protection, 7 had an app community, 9 allowed sharing, 9 required logins, and all 10 required web access.

There was no single dedicated focus described for any of the 10 apps, and there was an overlap between categories. Four of the apps had additional categories that were not prespecified (other), these linked directly to online stores to upgrade the app or buy baby-related products (4/10 apps). The entertainment category included apps that had links to online communities and information that was not always related to pregnancy health, and well-being. The physical health category included apps that provided information about pregnancy symptoms, milestones, nutrition, exercise, maternal weight gain, medication, and prenatal vitamin reminders (10/10 apps). Of the 10 apps, 5 had a component that related to supporting mental health (reducing negative feelings, anxiety, and depression).

All 10 apps provided some health information and education. All app descriptions stated that the app would help monitor and track various healthy behaviors, provide advice, tips, and strategies; 9 out of 10 apps had such content. Further analysis of the content using the MARS tool showed that few provided the necessary goal-setting (4/10 apps), assessment (3/10 apps), and feedback (2/10 apps) required to successfully support behavior change.

App Quality

The mean of the 4 MARS subscale scores across all 10 apps were 3.01 (range 1.97-4.40) in 2017 and 3.40 (range 2.27-4.44) in 2019. The app that had the highest quality scores was Ovia Pregnancy Tracker App (in 2017: mean 4.40; in 2019: mean 4.44). The app that had the lowest quality score in 2017 (mean 1.97) was Baby Bump Pregnancy Pro. This app was no longer available for download in 2019, therefore, was not analyzed at the second time point. In both 2017 and 2019, functionality was rated the highest in both 2017 and 2019 followed by aesthetics, engagement, and information (Table 1). Ovia Pregnancy Tracker scored highest across all subscales. Apps that scored higher for engagement and aesthetics scored lower for information. We found that sources of information were not cited, and studies and trials were not included. There was inconsistent information, and there appeared to be an ad hoc approach with several pieces of content missing review dates. The subjective quality items were calculated as a mean score. Apps that had the highest subjective scores in 2017 were Ovia Pregnancy Tracker (mean 3.75) and Sprout Pregnancy (mean 2.75). These 2 apps also had the highest overall MARS scores in 2017 (mean 4.40 and mean 3.38, respectively).

Table 1. MARS scores of the Top 10 apps in 2017 and 2019.

App	Engagement score		Functionality score		Aesthetics score		Information score		Overall quality score	
	2017	2019	2017	2019	2017	2019	2017	2019	2017	2019
All, mean	3.00	3.40	3.62	4.25	3.30	4.00	2.10	2.14	3.01	3.40
Ovia Pregnancy Tracker	4.60	4.60	5.00	4.75	5.00	5.00	3.00	3.42	4.40	4.44
I'm Expecting	3.00	2.20	3.50	3.00	3.00	2.66	2.11	2.00	2.90	2.46
Baby Centre	3.00	3.60	3.00	4.50	3.00	5.00	2.70	2.85	2.90	3.98
Pregnancy +	3.00	3.40	3.75	4.25	4.10	4.33	1.70	2.22	3.11	3.55
Glow	3.40	3.00	4.00	4.33	3.30	3.66	2.10	2.14	3.20	3.28
What to Expect	3.00	4.00	3.50	4.50	3.30	4.66	1.42	2.42	2.80	3.89
Baby Bump Pregnancy Pro	2.20	__[a]	2.20	—	2.30	—	1.20	—	1.97	—
Sprout Pregnancy	3.00	2.60	4.00	2.25	4.33	2.66	3.00	1.57	3.58	2.27
Week by Week	2.20	2.80	4.00	3.50	4.00	3.33	2.28	2.14	3.12	2.84
The Bump	2.00	4.20	2.75	4.00	3.00	4.00	1.40	1.42	2.28	3.40

[a]Denotes that this app was not available for download in 2019.

Comparative Analysis

Baby Bump Pregnancy Pro had the lowest MARS score in 2017 and was not available to download at the second time point on either iOS or Android. For the majority of apps, there was an increase in the mean quality scores from 2017 to 2019. The exceptions to this were I'm Expecting, Sprout Pregnancy, and Week by Week, which had lower scores than their 2017 scores. Sprout Pregnancy had the lowest star rating (4.5) at both time points. This app dropped significantly in ranking from 2017 (ranked second) to 2019 (ranked ninth). Ovia Pregnancy Tracker ranked first at both time points (Table 2).

Table 2. Comparison of Android downloads (rounded estimates as shown in Google Play) and star ratings for the top 10 apps.

App	Downloads		Star ratings	
	2017	2019	2017	2019
Ovia Pregnancy Tracker	81,053	1,000,000	4.8	4.8
I'm Expecting	70,640	1,000,000	4.6	4.6
Baby Centre	530,321	10,000,000	4.7	4.7
Pregnancy +	157,385	10,000,000	4.5	4.6
Glow	16,938	500,000	4.6	4.6
What to Expect	43,882	1,000,000	4.5	4.6
Baby Bump Pregnancy Pro	26,659	—[a]	4.5	—
Sprout Pregnancy	23,131	1,000,000	4.5	4.5
Week by Week	1100	1,000,000	4.8	4.9
The Bump	13,532	1,000,000	4.5	4.7

[a]Denotes that this app was not available for download in 2019.

Step 3: Analysis of Behavior Change Techniques Used

Overall, Ovia Pregnancy Tracker scored the highest for perceived impact using the MARS tool across all 3 behaviors (breastfeeding: mean 3.0; healthy weight: mean 3.5; maternal fetal movement awareness: mean 4.0; all apps: range 1-4). When examined using CALO-RE, Ovia Pregnancy Tracker did not have the highest number of behavior change techniques for any of the behaviors reviewed (Table 3). What to Expect used the highest number of behavior change techniques (breastfeeding: 11/40; healthy weight: 9/40; maternal fetal movement awareness: 8/40; all apps: range 2-12) (Table 3). A detailed analysis of the frequency of CALO-RE behavior change techniques included across the 10 apps for the 3 behaviors is included in Multimedia Appendix 1.

Fetal Movements

Ovia Pregnancy Tracker scored the highest for perceived impact for fetal movements (mean 4.0; range 1-4) (Table 3). When assessed using CALO-RE, the highest number of behavior change techniques used was 8/40 (range 2-8). When cross-referenced with the Australian Safer Baby Bundle Handbook and Resource Guide [39], we found several discrepancies. All apps had some inaccurate or incomplete information about maternal fetal movement monitoring. One app stated that normal baby movement was 10 kicks in 2 hours and all other apps provide no or partial information alongside the in-app tools provided. Three of the 10 apps incorrectly suggested the mother should consume something sweet to encourage fetal movements. One app recommended buying a fetal Doppler ultrasound, claiming that it was beneficial and so that the pregnant woman could monitor fetal well-being. This is not evidence-based and may impact negatively on fetal outcomes [42]. Three of the apps did not articulate or encourage women to contact a health care provider if concerned about decreased fetal movements or mention the risk of stillbirth.

Healthy Weight in Pregnancy

Ovia Pregnancy Tracker scored the highest for perceived impact (mean 3.5, range 1.0-3.5). Basic information on diet and exercise was included in all apps, with a focus primarily on fitness rather than weight management. Tools to track exercise and weight were included in 5 of the apps; however, little or no information was given on how, when, or why it is important to do so. Weight tracking tools in 2 apps provided incorrect information on expected weight gain, and 2 apps provided information that was misleading regarding increasing calorie intake and advocating the need to eat for two. CALO-RE analysis showed that Baby Centre (12 techniques) and What to Expect (9 techniques) utilized the highest of behavior change techniques (all apps: mean 4.4, range: 2-12), but there was no clear alignment with Institute of Medicine, National Institute for Health and Clinical Excellence, or Australian pregnancy care guidelines (Table 3).

Breastfeeding

Basic breastfeeding information was provided in all apps; however, the content did not adequately cover all aspects of breastfeeding, was inaccurate, or did not follow best practice outlined in the WHO code [41]. Although Ovia Pregnancy Tracker scored the highest for perceived impact (mean 3.0), it had one of the lowest numbers of behavior change techniques (5/40) when compared against other apps (What to Expect: 11/40; The Bump: 10/40). Two apps provided information for later in pregnancy and highlighted some difficulties with breastfeeding. It was noted that these apps did not mention midwives or lactation consultants as a form of support, instead suggesting that formula is equal to breastmilk. The apps varied greatly; 1 app provided a feeding tracker and gave links to relevant articles while another app gave information that was directly linked to online shopping for nipple creams and bras. One app had affiliations with a company that sells breastmilk substitutes; this app scored the lowest in both MARS and CALO-RE and directly contradicted the WHO code [41].

Table 3. Perceived impact of app on modifiable healthy behavior using the MARS app-specific tool and CALO-RE behavior change techniques taxonomy.

App	Breastfeeding		Healthy weight		Fetal movements	
	MARS[a] perceived impact score, mean	CALO-RE[b] behavior change techniques, n (%)	MARS perceived impact score, mean	CALO-RE behavior change techniques, n (%)	MARS perceived impact score, mean	CALO-RE behavior change techniques, n (%)
Ovia Pregnancy Tracker	3.0	5 (12.5)	3.5	4 (10)	4.0	4 (10)
I'm Expecting	1.1	2 (5)	1.6	2 (5)	1.6	5 (12.5)
Baby Centre	1.1	9 (22.5)	1.3	12 (30)	1.3	7 (17.5)
Pregnancy +	1.0	6 (15)	3.0	3 (7.5)	3.0	7 (17.5)
Glow	2.5	2 (5)	1.8	4 (10)	1.0	3 (7.5)
What to Expect	3.0	11 (27.5)	2.3	9 (22.5)	1.5	8 (20)
Baby Bump Pregnancy Pro	1.0	—[c]	1.3	—	1.3	—
Sprout Pregnancy	1.0	2 (5)	1.5	2 (5)	2.3	2 (5)
Week by Week	1.0	2 (5)	2.3	2 (5)	2.3	8 (20)
The Bump	1.0	10 (25)	1.0	2 (5)	1.0	8 (20)

[a]MARS: Mobile Application Rating Scale.

[b]CALO-RE: Coventry, Aberdeen, and London—Refined.

[c]Denotes that this app was not available for download in 2019.

Discussion

Principal Findings

This review showed that highly rated pregnancy smartphone apps were generally of low to moderate quality. This study is the first to have systematically described app quality and the use of behavior change techniques in pregnancy apps for 3 key health behaviors—maternal fetal movement monitoring, pregnancy weight monitoring, and breastfeeding.

The trend toward the use of mobile health opens up an opportunity to reach women who are less likely to or have yet to engage with health care providers. The information provided in all 10 apps, however, was not tailored for specific groups, for example, young mothers. We were unable to find any information in the 10 apps that suggested that pregnant women were engaged as co-designers at any stage of the app development or that endorsement from key maternity care organizations, health departments, or colleges was sought. All apps required web access. This may be a barrier for women living outside major Australian cities with limited Wi-Fi connection or with restricted data plans. Those who may benefit most are potentially not able to access interventions that may improve health outcomes and support pregnancy and early parenthood.

Functionality appears to be the main focus for developers. All 10 apps work technically well and have some in-built mechanisms for sharing, basic privacy settings (to meet Australian law), and rudimentary personalization capability. All reviewed apps lacked transparency regarding affiliations and have been set up to be commercial rather than as an intervention to change behavior. Although the functionality and usability of the apps have increased in the last two years, content credibility has not. Several reviews have found that health apps have insufficient evidence-based content [43-45]. Since there is limited regulatory oversight of the quality of apps and content provided, the MARS tool could be used by clinicians and by women themselves to identify high-quality apps rather than relying on download and star ratings. Our results confirm that few apps provide evidence-based information; therefore, caution is advised before recommending the use of these during pregnancy. For example, regarding maternal fetal movement awareness, reviewed app tools appeared to be for entertainment purposes since they were designed poorly and lacked essential information. We were unable to determine if fetal movement tools in apps during pregnancy would positively impact maternal knowledge, behavior change, or perinatal health outcomes. Of concern is the possibility of women assuming that app content is evidence-based and credible; however, in reality, apps are a platform for in-app purchases and link to unnecessary and potentially harmful advice. This was highlighted in the breastfeeding information with links to breastmilk substitutes and pictures of idealized artificial feeding. Such breaches of the WHO code do not contribute to motivating women to breastfeed exclusively for 6 months or seek help and assistance to do so.

Strengths and Limitations

The strength of this study is the use of a stepwise systematic approach to identify and review pregnancy mobile apps in Australia. By assessing current apps using the MARS tool and CALO-RE taxonomy, we were able to evaluate features and quality, as well as capture the usefulness of these apps as behavior change intervention strategies. We looked at the top 10 apps in terms of popularity using downloads and star ratings in Australian iOS and Android app stores. We then assessed 3

prespecified healthy pregnancy behaviors. Our inclusion criteria were deliberately set wide to assess what apps are commonly used by pregnant women; the disadvantage of using this criteria was that we may have missed other behaviors.

Our study supports the work by Brown et al who reviewed the quality of Australian iPhone pregnancy apps and the inclusion of behavior change techniques for pregnancy-specific nutrition information using MARS and CALO-RE [34] and free pregnancy apps available in the Google store [35]. Similarly, we found that only a small number of behavior change techniques were utilized across the 3 healthy behaviors we examined using CALO-RE. Likewise, our MARS findings had an overall mean quality score of 3.02 compared to the mean of 3.05 in Brown et al [34]. We found no single tool could cover all aspects of pregnancy apps and therefore we used both the MARS tool and the CALO-RE taxonomy. We chose the MARS tool as it has been validated for the quality of health apps. CALO-RE was used because it has more detail about specific behavior change techniques.

CALO-RE and perceived impact scores in this study are low. This may be attributed to the lack of behavior change theory in app design or that behavior change was not the purpose of the apps. The number of behavior change techniques used and the perceived impact does not necessarily correlate with actual behavior change as a result of using these apps. A feasibility study would need to be conducted to establish the link between perceived and actual impact, this is beyond the scope of this study; however, using CALO-RE and the MARS tool to assess the likelihood of behavior change has highlighted the need for improvement if apps are to be used as interventions.

Finally, searching for pregnancy apps is problematic due to inconsistent search terms across iTunes and Google Play. The implication of this, is that a search cannot be replicated, and therefore, validated. A potential solution to facilitate searching would be the development of a vocabulary for app indexing similar to Medical Subject Headings [46]. This would enable users and researchers alike, to have the ability to easily find the most appropriate apps with pregnancy information. Also, with the constant addition and removal of apps from the market, it is difficult to provide a timely appraisal of the current apps available. Future research could explore the creation of a combined scoring tool for pregnancy apps that could be used by both clinicians and women.

Conclusions

This study confirms that publicly available, free pregnancy apps in Australia should be treated with caution rather than recommended. Clinicians and researchers need to work collaboratively and show leadership in developing evidence-based pregnancy apps that incorporate behavior change techniques. Engagement with pregnant women in co-design must occur at all stages of app development, and endorsement from peak maternity care organizations, health departments, and professional societies should be sought. Smartphone apps have the potential to influence healthy behaviors in pregnancy, but an evidence-based approach is needed.

Acknowledgments

This study is part of LM's doctoral thesis, which is funded through the Ho Kong Fung Ling postgraduate scholarship.

Authors' Contributions

This project was conducted as part of a Doctor of Philosophy. LM and AG contributed to the concepts and design of the study. LM and AG conducted the research and analyzed the data. LM drafted the first version of the manuscript. AG, NK, and CH contributed to writing and editing the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Frequency of behavior change technique inclusion.
[DOCX File , 35 KB - mhealth_v8i11e22340_app1.docx]

References

1. Silver L. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. Pew Research Center. 2019 Feb 5. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2020-05-07]
2. Communications Report 2018-19. Australian Communications and Media Authority. Australia URL: https://www.acma.gov.au/sites/default/files/2020-04/Communications%20report%202018-19.pdf [accessed 2020-05-07]
3. Household Use of Information Technology. Australian Bureau of Statistics. Australia; 2018. URL: https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/8146.0Main+Features12016-17?OpenDocument [accessed 2020-05-07]
4. Mobile Consumer Survey 2019. Deloitte Australia. Australia; 2019. URL: http://images.content.deloitte.com.au/Web/DELOITTEAUSTRALIA/

%7B3a35a2e7-7ce2-4cc9-b7b2-3472be4267a7%7D_Deloitte_Global_Mobile_Consumer_Survey_2019.pdf [accessed 2020-05-07]
5. Kennedy R, Mullaney L, Reynolds C, Cawley S, McCartney D, Turner M. Preferences of women for web-based nutritional information in pregnancy. Public Health 2017 Feb;143:71-77. [doi: 10.1016/j.puhe.2016.10.028] [Medline: 28159029]
6. Rodger D, Skuse A, Wilmore M, Humphreys S, Dalton J, Flabouris M, et al. Pregnant women's use of information and communications technologies to access pregnancy-related health information in South Australia. Aust J Prim Health 2013;19(4):308. [doi: 10.1071/py13029]
7. Kraschnewski J, Chuang C, Poole E, Peyton T, Blubaugh I, Pauli J, et al. Paging "Dr. Google": does technology fill the gap created by the prenatal care visit structure? Qualitative focus group study with pregnant women. J Med Internet Res 2014 Jun 03;16(6):e147 [FREE Full text] [doi: 10.2196/jmir.3385] [Medline: 24892583]
8. O'Higgins A, Murphy OC, Egan A, Mullaney L, Sheehan S, Turner MJ. The use of digital media by women using the maternity services in a developed country. Ir Med J 2014;107(10):313-315. [Medline: 25556255]
9. Hearn L, Miller M, Lester L. Reaching perinatal women online: the Healthy You, Healthy Baby website and app. J Obes 2014;2014:573928 [FREE Full text] [doi: 10.1155/2014/573928] [Medline: 24872891]
10. Hearn L, Miller M, Fletcher A. Online healthy lifestyle support in the perinatal period: what do women want and do they use it? Aust J Prim Health 2013;19(4):313. [doi: 10.1071/py13039]
11. Walker LO, Mackert MS, Ahn J, Vaughan MW, Sterling BS, Guy S, et al. e-Health and new moms: Contextual factors associated with sources of health information. Public Health Nurs 2017 Nov;34(6):561-568. [doi: 10.1111/phn.12347] [Medline: 28762533]
12. Phelan S. Pregnancy: a "teachable moment" for weight control and obesity prevention. Am J Obstet Gynecol 2010 Feb;202(2):135.e1-135.e8 [FREE Full text] [doi: 10.1016/j.ajog.2009.06.008] [Medline: 19683692]
13. Olander EK, Darwin ZJ, Atkinson L, Smith DM, Gardner B. Beyond the 'teachable moment' - A conceptual analysis of women's perinatal behaviour change. Women Birth 2016 Jun;29(3):e67-e71. [doi: 10.1016/j.wombi.2015.11.005] [Medline: 26626592]
14. Ashman A, Collins C, Brown L, Rae K, Rollo M. Validation of a Smartphone Image-Based Dietary Assessment Method for Pregnant Women. Nutrients 2017 Jan 18;9(1) [FREE Full text] [doi: 10.3390/nu9010073] [Medline: 28106758]
15. Wang N, Deng Z, Wen LM, Ding Y, He G. Understanding the Use of Smartphone Apps for Health Information Among Pregnant Chinese Women: Mixed Methods Study. JMIR Mhealth Uhealth 2019 Jun 18;7(6):e12631 [FREE Full text] [doi: 10.2196/12631] [Medline: 31215516]
16. Asiodu IV, Waters CM, Dailey DE, Lee KA, Lyndon A. Breastfeeding and use of social media among first-time African American mothers. J Obstet Gynecol Neonatal Nurs 2015;44(2):268-278 [FREE Full text] [doi: 10.1111/1552-6909.12552] [Medline: 25712127]
17. Olson CM, Groth SW, Graham ML, Reschke JE, Strawderman MS, Fernandez ID. The effectiveness of an online intervention in preventing excessive gestational weight gain: the e-moms roc randomized controlled trial. BMC Pregnancy Childbirth 2018 May 09;18(1):148 [FREE Full text] [doi: 10.1186/s12884-018-1767-4] [Medline: 29743026]
18. Daly LM, Horey D, Middleton PF, Boyle FM, Flenady V. The Effect of Mobile App Interventions on Influencing Healthy Maternal Behavior and Improving Perinatal Health Outcomes: Systematic Review. JMIR Mhealth Uhealth 2018 Aug 09;6(8):e10012 [FREE Full text] [doi: 10.2196/10012] [Medline: 30093368]
19. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The Usability and Effectiveness of Mobile Health Technology-Based Lifestyle and Medical Intervention Apps Supporting Health Care During Pregnancy: Systematic Review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
20. Russell CG, Denney-Wilson E, Laws RA, Abbott G, Zheng M, Lymer SJ, et al. Impact of the Growing Healthy mHealth Program on Maternal Feeding Practices, Infant Food Preferences, and Satiety Responsiveness: Quasi-Experimental Study. JMIR Mhealth Uhealth 2018 Apr 25;6(4):e77 [FREE Full text] [doi: 10.2196/mhealth.9303] [Medline: 29695373]
21. Flenady V, Gardener G, Boyle FM, Callander E, Coory M, East C, IMPACT Clinical Trials Network for Mothers'Babies' Health. My Baby's Movements: a stepped wedge cluster randomised controlled trial to raise maternal awareness of fetal movements during pregnancy study protocol. BMC Pregnancy Childbirth 2019 Nov 21;19(1):430 [FREE Full text] [doi: 10.1186/s12884-019-2575-1] [Medline: 31752771]
22. Goldstein RF, Abell SK, Ranasinha S, Misso M, Boyle JA, Black MH, et al. Association of Gestational Weight Gain With Maternal and Infant Outcomes: A Systematic Review and Meta-analysis. JAMA 2017 Jun 06;317(21):2207-2225 [FREE Full text] [doi: 10.1001/jama.2017.3635] [Medline: 28586887]
23. Thangaratinam S, Rogozinska E, Jolly K, Glinkowski S, Roseboom T, Tomlinson JW, et al. Effects of interventions in pregnancy on maternal weight and obstetric outcomes: meta-analysis of randomised evidence. BMJ 2012 May 16;344:e2088 [FREE Full text] [doi: 10.1136/bmj.e2088] [Medline: 22596383]
24. Institute of Medicine. Weight Gain During Pregnancy: Reexamining the Guidelines. USA: National Academies Press; 2009. URL: https://books.google.com.au/books?id=jAAwbDUnH4MC [accessed 2020-05-08]

25. National Institute for Health and Clinical Excellence. Antenatal care for uncomplicated pregnancies. National Institute for Health and Clinical Excellence. UK: National Institute for Health and Clinical Excellence; 2019 Feb 04. URL: https://www.nice.org.uk/guidance/cg62/resources/antenatal-care-for-uncomplicated-pregnancies-pdf-975564597445 [accessed 2020-05-10]
26. Clinical Practice Guidelines: Pregnancy Care. Australia: Australian Government Department of Health; 2018. URL: https://www.health.gov.au/resources/publications/pregnancy-care-guidelines [accessed 2020-05-08]
27. Bland C, Dalrymple KV, White SL, Moore A, Poston L, Flynn AC. Smartphone applications available to pregnant women in the United Kingdom: An assessment of nutritional information. Matern Child Nutr 2020 Apr;16(2):e12918 [FREE Full text] [doi: 10.1111/mcn.12918] [Medline: 31833237]
28. Victora CG, Bahl R, Barros AJD, França GVA, Horton S, Krasevec J, et al. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. The Lancet 2016 Jan;387(10017):475-490. [doi: 10.1016/s0140-6736(15)01024-7]
29. Tang K, Gerling K, Chen W, Geurts L. Information and Communication Systems to Tackle Barriers to Breastfeeding: Systematic Search and Review. J Med Internet Res 2019 Sep 27;21(9):e13947 [FREE Full text] [doi: 10.2196/13947] [Medline: 31573903]
30. Deave T, Ginja S, Goodenough T, Bailey E, Piwek L, Coad J, et al. The Bumps and BaBies Longitudinal Study (BaBBLeS): a multi-site cohort study of first-time mothers to evaluate the effectiveness of the Baby Buddy app. Mhealth 2019;5:42 [FREE Full text] [doi: 10.21037/mhealth.2019.08.05] [Medline: 31620469]
31. 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]
32. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
33. 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]
34. Brown HM, Bucher T, Collins CE, Rollo ME. A review of pregnancy iPhone apps assessing their quality, inclusion of behaviour change techniques, and nutrition information. Matern Child Nutr 2019 Jul;15(3):e12768. [doi: 10.1111/mcn.12768] [Medline: 30569549]
35. Brown HM, Bucher T, Collins CE, Rollo ME. A review of pregnancy apps freely available in the Google Play Store. Health Promot J Austr 2020 Sep;31(3):340-342. [doi: 10.1002/hpja.270] [Medline: 31225924]
36. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile Phone Apps to Improve Medication Adherence: A Systematic Stepwise Process to Identify High-Quality Apps. JMIR Mhealth Uhealth 2016 Dec 02;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
37. Furlong L, Morris M, Serry T, Erickson S. Mobile apps for treatment of speech disorders in children: An evidence-based analysis of quality and efficacy. PLoS One 2018;13(8):e0201513 [FREE Full text] [doi: 10.1371/journal.pone.0201513] [Medline: 30092095]
38. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the Quality of Mobile Apps for the Management of Pain: Systematic Search and Evaluation Using the Mobile App Rating Scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718 [FREE Full text] [doi: 10.2196/10718] [Medline: 30361196]
39. Safer Baby Bundle Handbook and Resource Guide: Working together to reduce stillbirth. Centre of Research Excellence Stillbirth. Australia; 2019. URL: https://resources.stillbirthcre.org.au/downloads/SBB+Handbook_Final.pdf [accessed 2020-05-08]
40. Protecting, promoting and supporting breastfeeding in facilities providing maternity and newborn services: implementing the revised Baby-friendly Hospital Initiative 2018. Geneva: World Health Organization and the United Nations Children's Fund (UNICEF); 2018. URL: https://www.who.int/nutrition/publications/infantfeeding/bfhi-implementation-2018.pdf [accessed 2020-05-08]
41. International code of marketing of breast-milk substitutes. Geneva: World Health Organization; 1981. URL: https://www.who.int/nutrition/publications/code_english.pdf [accessed 2020-05-10]
42. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? Women Birth 2014 Mar;27(1):64-67. [doi: 10.1016/j.wombi.2013.11.001] [Medline: 24295598]
43. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
44. 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?: a content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]
45. Rivera J, McPherson A, Hamilton J, Birken C, Coons M, Iyer S, et al. Mobile Apps for Weight Management: A Scoping Review. JMIR Mhealth Uhealth 2016 Jul 26;4(3):e87 [FREE Full text] [doi: 10.2196/mhealth.5115] [Medline: 27460502]

46. Medical Subject Headings. USA: National Library of Medicine; 2020. URL: https://www.nlm.nih.gov/mesh/meshhome.html [accessed 2020-05-26]

Abbreviations

CALO-RE: Coventry, Aberdeen, and London—Refined
MARS: Mobile Application Rating Scale
mHealth: mobile health
WHO: World Health Organization

Edited by G Eysenbach, L Buis; submitted 09.07.20; peer-reviewed by H Cheng, E Neter; comments to author 17.08.20; revised version received 22.09.20; accepted 26.10.20; published 16.11.20.

Please cite as:
Musgrave LM, Kizirian NV, Homer CSE, Gordon A
Mobile Phone Apps in Australia for Improving Pregnancy Outcomes: Systematic Search on App Stores
JMIR Mhealth Uhealth 2020;8(11):e22340
URL: http://mhealth.jmir.org/2020/11/e22340/
doi:10.2196/22340
PMID:33196454

Original Paper

Comparison of Mobile Health Technology Use for Self-Tracking Between Older Adults and the General Adult Population in Canada: Cross-Sectional Survey

Mirou Jaana[1], PhD; Guy Paré[2], PhD

[1]Telfer School of Management, University of Ottawa, Ottawa, ON, Canada

[2]Research Chair in Digital Health, HEC Montreal, Montreal, QC, Canada

Corresponding Author:
Mirou Jaana, PhD
Telfer School of Management
University of Ottawa
55 Laurier Ave East
Ottawa, ON, K1N 6N5
Canada
Phone: 1 16135625800 ext 3400
Email: jaana@telfer.uottawa.ca

Abstract

Background: The burden of population aging and chronic conditions has been reported worldwide. Older adults, especially those with high needs, experience social isolation and have high rates of emergency visits and limited satisfaction with the care they receive. Mobile health (mHealth) technologies present opportunities to address these challenges. To date, limited information is available on Canadian older adults' attitudes toward and use of mHealth technologies for self-tracking purposes—an area that is increasingly important and relevant during the COVID-19 era.

Objective: This study presents contributions to an underresearched area on older adults and mHealth technology use. The aim of this study was to compare older adults' use of mHealth technologies to that of the general adult population in Canada and to investigate the factors that affect their use.

Methods: A cross-sectional survey on mHealth and digital self-tracking was conducted. A web-based questionnaire was administered to a national sample of 4109 Canadian residents who spoke either English or French. The survey instrument consisted of 3 sections assessing the following items: (1) demographic characteristics, health status, and comorbidities; (2) familiarity with and use of mHealth technologies (ie, mobile apps, consumer smart devices/wearables such as vital signs monitors, bathroom scales, fitness trackers, intelligent clothing); and (3) factors influencing the continued use of mHealth technologies.

Results: Significant differences were observed between the older adults and the general adult population in the use of smart technologies and internet (P<.001). Approximately 47.4% (323/682) of the older adults in the community reported using smartphones and 49.8% (340/682) indicated using digital tablets. Only 19.6% (91/463) of the older adults using smartphones/digital tablets reported downloading mobile apps, and 12.3% (47/383) of the older adults who heard of smart devices/wearables indicated using them. The majority of the mobile apps downloaded by older adults was health-related; interestingly, their use was sustained over a longer period of time (P=.007) by the older adults compared to that by the general population. Approximately 62.7% (428/682) of the older adults reported tracking their health measures, but the majority did so manually. Older adults with one or more chronic conditions were mostly nontrackers (odds ratio 0.439 and 0.431 for traditional trackers and digital trackers, respectively). No significant differences were observed between the older adults and the general adult population with regard to satisfaction with mHealth technologies and their intention to continue using them.

Conclusions: Leveraging mHealth technologies in partnership with health care providers and sharing of health/well-being data with health care professionals and family members remain very limited. A culture shift in the provision of care to older adults is deemed necessary to keep up with the development of mHealth technologies and the changing demographics and expectations of patients and their caregivers.

(JMIR Mhealth Uhealth 2020;8(11):e24718) doi:10.2196/24718

KEYWORDS

mobile health; older adults; self-tracking; wearable technology; smart devices; mobile apps; survey; mobile phone; seniors; elderly

Introduction

Population aging is a phenomenon that is associated with increased prevalence of chronic conditions worldwide [1-3]. In 2017, the global number of people 60 years and older was 962 million (including 137 million ≥80 years), and this number is expected to reach 2.1 billion by 2050 [3]. This growing population of older adults leads to an increased demand on the health systems for services, which are costly and require significant resources [4].

A recent Commonwealth survey of older adults in 11 countries investigated the challenges faced by adults aged ≥65 years at the social and health care levels [2,5]. The results showed that, across all surveyed countries, older adults, in general, and older adults with high needs, in particular (ie, multiple chronic conditions/functional challenges), experience social isolation and have high rates of emergency visits and general dissatisfaction with the quality of care they receive [2]. The current COVID-19 crisis has further catalyzed this problem. Older adults represent a group of the population that is at higher risk of death from severe acute respiratory syndrome associated with coronavirus, thus necessitating social distancing, which may lead to social isolation [6]. This confinement and social isolation can in turn have negative psychological effects and sleeping problems [7,8] and increased risk for early mortality [9]. In addition to potential social isolation, COVID-19 may have long-term effects on people with preexisting noncommunicable chronic diseases [10]. This is particularly observed with decrease in physical activity, unhealthy lifestyles during the COVID-19 crisis, and changes in the management of these conditions (eg, reduced outpatient visits, difficulty in diagnosing new conditions or recognizing deterioration in the existing ones) [10].

In Canada, wait times for various types of services (eg, doctors, specialists, emergency) have been historically longer compared to some other developed countries [11]. Generally, Canadians tend to be more frequent users of health services [11], and concerns have been growing about the ability of the public health care system to address the increasing needs of an aging population [12]. For instance, the population of Canadian adults aged 65 years and older reached 6.5 million in 2019 [13], and this number is expected to increase by 68% over the next 20 years [14]. This is particularly critical, given the high health care spending per capita on older adults and large use of services by this group [12].

Mobile health (mHealth) technologies present an opportunity to address the challenges associated with population aging and enable support for older adults in the community. mHealth refers to the use of mobile devices (eg, patient monitoring devices, mobile phones) to detect and monitor physiological changes and support medical and public health practice [15]. Prior research has examined the potential role of mHealth technologies in providing long-term support for older adults [16-18] and in monitoring chronic conditions often associated with older age [19-26]. Self-tracking devices in particular (eg, smart devices with mobile apps, fitness trackers, blood pressure monitors) have gained interest in recent years in light of their potential for monitoring and motivating individuals to remain healthy [27-31]. However, their use remains variable and less widespread among older adults [32], and prior research has reported risks associated with health information tracking, which may trigger negative emotions among patients with multiple chronic conditions and potential emotional draining in this group [33]. With the current COVID-19 crisis, calls for initiatives and efforts to bridge health information and communication technologies with the care for older adults have appeared in various countries as a preparedness mechanism and a mitigating approach against the current and future pandemics [7,10,34,35]. However, to date, limited information is available on older adults' attitudes toward and their use of mHealth technologies for self-tracking purposes—an area that reveals to be increasingly important during and following this COVID-19 era.

This study, which is part of a larger program on digital health self-tracking [36], addresses this gap and presents the results of a national survey across all provinces in Canada, which assessed older adults' familiarity with and use of mHealth technologies comprising mobile apps, smart devices, and wearables. Specifically, we report findings on the pattern of older adults' use of mHealth technologies for self-tracking purposes and compare it to that of the general adult population. We also investigate the factors that influence the continued usage of mHealth technologies among older adults.

In order to address the objectives of this study, we propose a research model based on the work of Bhattacherjee [37] and Hong et al [38] and the expectation-confirmation theory [39]. In the present context, this model suggests that an older adult's intention to continue using mHealth technologies is mainly influenced by his or her level of satisfaction, which is in turn affected by the extent to which his or her initial expectations toward mHealth technologies are confirmed, in addition to ease of use and perceived usefulness [40]. The latter factors also have direct links with usage continuance intention [38]. Hence, this study presents evidence on the extent to which older adults use mHealth technologies to self-track their health, compares their use of mHealth technologies to that of the general adult population, and analyzes the factors that influence the continued use of these technologies in the older population.

Methods

Study Design and Sampling

We present in this section the survey that was conducted in accordance with the Checklist for Reporting Results of Internet E-Surveys [41]. A web-based questionnaire was administered to a national sample of 4109 Canadian residents, aged 18 years or older, and who spoke English or French. The sample was selected from a proprietary web-based panel (AC Nielsen Company), which is one of the largest and most representative

panels in Canada. To ensure representativeness of the overall population, the quota method was applied (age and gender) following a stratification by the geographic region. The ethics approval for the study was granted by the HEC Montréal's research ethics committee. The older adult group in the sample consisted of all respondents aged 65 years and older and the general adult population in the sample consisted of respondents aged 18-64 years.

Survey and Data Collection

The survey instrument consisted of 3 sections assessing the following items: (1) demographic characteristics, health status, and comorbidities; (2) familiarity with and use of mHealth technologies (ie, mobile apps, consumer smart devices/wearables such as vital signs monitors, bathroom scales, fitness trackers, intelligent clothing); and (3) factors influencing the continued use of mHealth technologies. The latter section also measured respondents' satisfaction, ease of use, expectation confirmation, perceived usefulness, and intention to continue using mHealth technologies in the future.

Sociodemographic variables were measured using standardized indicators used in other international surveys [42-45]. These included gender, age, region, gross family income, education, occupation, and use of mobile phones and digital tablets. Health status was self-rated by respondents on a 5-point Likert scale (1=poor or fair, 5=very good or excellent), which is a common approach used in prior research [46]. A total of 11 chronic conditions were investigated (eg, diabetes, high blood pressure, cardiovascular disease, lung or respiratory cancer).

Respondents' familiarity with mHealth technologies was assessed using a combination of items. A general question measured their familiarity with consumer wearables and smart medical devices on a 5-point Likert scale (1=not much at all familiar, 5=extremely familiar). Participants were also asked to indicate the devices that they owned by using descriptive terms that referred to 13 devices identified in the literature and available in the Canadian market. When participants indicated owning a specific device or wearable, they were asked to rate on a (1-7) scale (1=once a month or less, 7=many times a day) how often they used it in the past 3 months.

Three self-tracking profiles were identified in this study based on the respondents' indication of their health tracking behavior. Those who regularly tracked one or more aspects of their health or well-being by using mHealth technologies, including mobile apps for health, consumer wearables (eg, fitness trackers), and smart medical devices (eg, blood pressure monitors), were defined as "digital trackers." Respondents who regularly monitored one or more aspects of their health and well-being by using manual tools (ie, recording the information in writing) were defined as "traditional trackers." All other respondents who did not regularly monitor any aspect of their personal health were considered as "nontrackers."

The factors that are likely to influence the continued usage of mHealth technologies were captured. First, measures of perceived usefulness (7 items) and ease of use (4 items) were adapted from Davis [40] and rated on a 5-point Likert scale (1=strongly disagree, 5=strongly agree). We also adapted measures from Bhattacherjee [37] and Hong et al [38] to assess users' satisfaction (3 items) and confirmation of initial expectations (3 items) on a 5-point Likert scale (1=strongly disagree, 5=strongly agree).

The survey instrument was first pretested during face-to-face interviews with 16 adults who were representative of the Canadian population in terms of gender, language, and age. Some small adjustments were made to the questionnaire following this step. A copy of the final survey instrument may be obtained from the authors upon request. Panel members were invited to participate in the study by email. Once participants clicked on the URL provided in the email letter, they were screened for the abovementioned eligibility criteria. All respondents read and approved an informed consent form prior to completing the questionnaire. Survey respondents were able to enter the survey at any point during the data collection period, ie, from January 11, 2017 to February 2, 2017. In accessing the web-based questionnaire, respondents were assigned a unique identifier and secret code (closed survey) that allowed them access to their data until the survey was done. Those who partly completed the survey were able to exit the questionnaire and return at a later time to enter additional data and to review and change their prior answers. Participants were rewarded gift cards (eg, Amazon, iTunes, Starbucks, magazine subscriptions) for survey completion. Rewards ranged in value from CAD $5 to CAD $75.

Statistical Analysis

Data analysis was conducted to explore and better understand the pattern of use of these technologies and self-tracking behaviors by older adults in the community and compare it to that of the general adult population. Descriptive data analysis was performed to present an overview of the older adult group characteristics and their use of mHealth technologies. Bivariate analyses (two-sided *t* test for continuous variables and chi-square for categorical variables) were conducted to assess the differences between the 2 groups on these variables. Multinomial logistic regression tests were used to compare self-trackers (traditional and digital) and nontrackers, and Pearson correlation tests and partial least squares multiple regression analyses were used to analyze users' appreciation of digital self-tracking devices. Data analyses were performed on SPSS Statistics v25 (IBM Corp) and SmartPLS 2.0 (SmartPLS GmbH).

Results

Sample Characteristics

Of the total study population of 4109 participants distributed across all provinces, 682 (16.6%) were aged 65 years and older (older adults) and 3427 (83.4%) were aged 18-64 years (general adult population), which represents the actual distribution of the older adults in the Canadian population [13]. Table 1 shows that a higher proportion of the older adults live on the east coast of Canada and British Columbia.

Table 1. Comparison of the characteristics of the older adults with those of the general adult population in this study.

Characteristics	Older adult population, n=682, n (%)	General population, n=3427, n (%)	Total, N=4109, n (%)	*P* value
Gender				<.001
Male	400 (58.6)	1718 (50.1)	2118 (51.5)	
Female	282 (41.3)	1709 (49.9)	1991 (48.5)	
Region[a]				<.001
Atlantic provinces	56 (8.2)	237 (6.9)	293 (7.1)	
Quebec	153 (22.4)	833 (24.3)	986 (24.0)	
Ontario	265 (38.9)	1310 (38.2)	1575 (38.3)	
Prairies	37 (5.4)	229 (6.7)	266 (6.5)	
Alberta	50 (7.3)	387 (11.3)	437 (10.6)	
British Columbia and territories	121 (17.7)	431 (12.6)	552 (13.4)	
Highest education level[b,c]				.09
Primary and secondary school	181 (26.7)	758 (22.5)	939 (23.2)	
College/CEGEP	177 (26.1)	972 (28.8)	1149 (28.4)	
University undergraduate	207 (30.6)	1093 (32.4)	1300 (32.1)	
University graduate	112 (16.5)	549 (16.3)	660 (16.3)	
Employment				<.001
Full-time	37 (5.4)	1921 (56.1)	1958 (47.6)	
Part-time	44 (6.4)	385 (11.2)	429 (10.4)	
Retired	587 (86.1)	350 (10.2)	937 (22.8)	
Other	14 (2.1)	771 (22.5)	785 (19.1)	
Income[d]				<.001
<$20,000	32 (5.7)	236 (8.1)	268 (7.7)	
$20,000-$39,999	123 (22.1)	461 (15.7)	584 (16.7)	
$40,000-$59,999	131 (23.5)	482 (16.4)	613 (17.6)	
$60,000-$79,999	95 (17.1)	465 (15.9)	560 (16.1)	
$80,000-$99,000	74 (13.2)	424 (14.4)	498 (14.3)	
≥$100,000	102 (18.3)	863 (29.4)	965 (27.7)	
Chronic conditions[e]				<.001
Yes	342 (51.4)	939 (28.0)	1281 (31.9)	
No	323 (48.6)	2413 (72.0)	2736 (68.1)	
Current health status[b]				.87
Very poor/poor	63 (9.2)	339 (9.9)	402 (9.8)	
Good	345 (50.7)	1724 (50.3)	2070 (50.4)	
Very good/excellent	274 (40.2)	1364 (39.8)	1638 (39.9)	
Tracking health measures				<.001
Manual self-tracking	307 (45.0)	744 (21.7)	1051 (25.6)	
Electronic self-tracking	121 (17.8)	1547 (45.1)	1668 (40.6)	
No self-tracking	254 (37.2)	1135 (33.1)	1389 (33.8)	

[a]Atlantic provinces include Nova Scotia, New Brunswick, New Foundland and Labrador, and Prince Edward Island; Prairies include Manitoba and Saskatchewan; Territories include Nunavut, Yukon, and Northwest territories.

[b]Significant differences were observed between seniors and the general population on all variables except for "Current health status" and the "Highest level of education".

[c]There were 5 and 55 nonrespondent in the older adult group and the general adult population, respectively.

[d]There were 125 and 496 nonrespondents in the older adult group and the general adult population, respectively. All income data are provided in Canadian dollars (CAD $1=US $1.31).

[e]There were 17 and 75 nonrespondents in the older adult group and the general adult population, respectively.

When comparing the older adult population with the general adult population, significant differences were observed in all the characteristics, with the exception of education level and reported health status; comparable educational levels were noted in the 2 groups and the perceived health status was reported as good-to-excellent in both groups. Compared to the general adult population, the older adult population had a larger number of men, who were retired, and had an annual income below CAD $60K. Of the 682 older adults, 342 (50.1%) indicated having one or more chronic conditions compared to 939 (27.4%) respondents of the total general adult population of 3427. Among the 62.8% (428/682) of the older adults who reported self-tracking of their health, 17.7% (121/682) did so electronically (digital trackers) compared to 45.1% (1547/3426) in the general adult population. The majority of the older adults reported tracking their health parameters manually (traditional trackers).

Internet and Smart Technologies

Table 2 shows significant differences between the older adult population and the general adult population in terms of internet and smart technology use. Of the 682 older adults, 323 (47.3%) and 340 (49.8%) reported using a smartphone and a digital tablet, respectively, as compared to 2887 (84.2%) and 2337 (68.2%) respondents of the 3427 respondents in the general adult population. Among the 463 older adults using smartphones/digital tablets (out of 682 participating in this study), only 91 respondents (19.6%) downloaded ≥1 mobile apps and 314 (67.8%) indicated accessing the internet on a daily basis versus 45.6% (1406/3082) and 87.9% (2709/3082) in the general adult population, respectively. When asked about their familiarity with smart devices/wearables for health, 82.7% of the older respondents (383/463) indicated having heard of these technologies, but only 32.1% of the older adults (123/383) were somewhat familiar or very familiar with them.

Table 2. Comparison of the internet and mobile health technology use of the older adults with that of the general adult population.

Use of internet and mobile health technology	Older population	General population	Total	*P* value
Using a smartphone	682	3427	4109	<.001
Yes, n (%)	323 (47.4)	2887 (84.2)	3210 (78.1)	
No, n (%)	359 (52.6)	540 (15.8)	899 (21.9)	
Using a digital tablet	682	3426	4109	<.001
Yes, n (%)	340 (49.9)	1997 (58.3)	2337 (56.9)	
No, n (%)	342 (50.1)	1429 (41.7)	1772 (43.1)	
Accessing internet using smartphone/digital tablet[a]	463	3082	3545	<.001
Never, n (%)	52 (11.2)	89 (2.9)	141 (4.0)	
Less than daily, n (%)	97 (21.0)	284 (9.2)	381 (10.7)	
Daily, n (%)	314 (67.8)	2709 (87.9)	3023 (85.3)	
Downloaded ≥1 mobile apps on smartphone/digital tablet[a]	463	3082	3545	<.001
Yes, n (%)	91 (19.6)	1406 (45.6)	1497 (42.2)	
No, n (%)	372 (80.3)	1676 (54.4)	2048 (57.8)	
Heard of smart devices/wearables for health	463	3082	3545	.03
Yes, n (%)	383 (82.7)	2667 (86.5)	3050 (86.0)	
No, n (%)	80 (17.3)	415 (13.5)	495 (14.0)	
Familiarity with smart devices/wearables for health[a]	383	2667	3050	<.001
Slightly familiar, n (%)	260 (67.9)	1227 (46.0)	1487 (48.7)	
Somewhat familiar, n (%)	103 (26.9)	973 (36.5)	1076 (35.3)	
Very familiar, n (%)	20 (5.2)	467 (17.5)	487 (16.0)	

[a]The total values in the rows indicate the number of respondents for that category, which may be lower than the total number of older adults and the general adult population.

Mobile Apps for Health and Well-being

Table 3 compares the use of mobile apps for health/well-being between the older adult population and the general adult population. Among the 91 older adults who downloaded ≥1 mobile apps (presented in Table 2), 78 respondents (86%) indicated having used mobile apps for health/well-being in the last 3 months, which is comparable to that of the general population (1257/1406 respondents, ie, 89.4%). No significant differences were noted in relation to the number of mobile apps for health used nor in the extent of data sharing between the 2 groups. Among the 30 older adults that reported sharing data, 21 respondents (70%) indicated sharing data with family members, and 4 respondents (13%) reported sharing data with their friends and doctors. Interestingly, 38% of the older adults (29/77) reported using these mobile apps for 1-2 years as compared to 22.1% in the general population (269/1219). However, it is important to note that no significant differences were observed between the older adults and the general population who used mobile apps for health in the factors that affect their use (ie, perceived ease of use, perceived usefulness, and expectation confirmation). The overall satisfaction and intention to continue using mobile apps were favorable in both groups.

Table 3. Comparison of the use and perceptions of mobile apps for health between older adults who indicated downloading these apps and their counterparts in the general adult population.

Use and perceptions	Older population	General population	Total	*P* value
Mobiles apps for health/well-being used (last 3 months)[a]	78	1257	1335	.06
1 app, n (%)	40 (51)	514 (40.9)	554 (41.5)	
2 apps, n (%)	26 (33)	406 (32.3)	432 (32.4)	
≥3 apps, n (%)	12 (15)	337 (26.8)	349 (26.1)	
Duration of use of mobile health/well-being apps[a]	77	1219	1296	.007
<1 year, n (%)	39 (51)	790 (64.8)	829 (64.0)	
1-2 years, n (%)	29 (38)	269 (22.1)	298 (23.0)	
>2 years, n (%)	9 (12)	160 (13.1)	169 (13.0)	
Sharing of health/well-being data from apps[a]	77	1238	1315	.50
Yes, n (%)	30 (39)	436 (35.2)	466 (35.4)	
No, n (%)	47 (61)	802 (64.8)	849 (64.6)	
Satisfaction with mobile apps, mean (min-max)[b]	3.79 (1.67-5)	3.78 (1-5)	3.70 (1-5)	.89
Ease of use, mean (min-max)[b]	4.00 (1.5-5)	3.95 (1-5)	3.95 (1-5)	.55
Expectation confirmation, mean (min-max)[b]	3.74 (1.67-5)	3.60 (1-5)	3.61 (1-5)	.12
Perceived usefulness, mean (min-max)[b]	3.59 (1.25-5)	3.56 (1-5)	3.56 (1-5)	.78
Intention to continue using mobile apps, mean (min-max)[b]	3.97 (1-5)	3.91 (1-5)	3.92 (1-5)	.61

[a]The total values in the rows indicate the number of respondents for that category, which may be lower than the total number of older adults and the general adult population.

[b]The means represent the average of 4 questions that constitute each scale (satisfaction with mobile apps, ease of use, expectation confirmation, perceived usefulness, and intention to continue using mobile apps). Continuous variables were measured on a 5-point Likert scale.

Smart Devices/Wearables for Health

Among the 383 older adults in the sample who had heard of smart devices/wearables (as presented in Table 2), 47 respondents (12.2%) reported having ≥1 smart devices and indicated currently using them, while another 35 respondents (9.1%) reported having these devices but not using them; the remaining 78.6% (302/383) indicated not having smart devices/wearables. The majority of the older adults had only 1 device as opposed to the general adult population with more respondents indicating having 2 or more devices (see Table 4), and the most common type of smart devices/wearables used was bracelet/wristband watches.

Table 4. Comparison of the use of smart devices/wearables for health between older adults who own these devices and their counterparts in the general adult population.

Use and perceptions	Older population	General population	Total	*P* value
Having ≥1 smart device/wearables for health, n (%)[a]	384	2667	3051	<.001
Yes, and use them	47 (12.2)	533 (20.0)	580 (19.0)	
Yes, and stopped using them	24 (6.3)	236 (8.9)	260 (8.5)	
Yes, and never used them	11 (2.8)	164 (6.1)	175 (5.7)	
No	302 (78.6)	1734 (65.0)	2036 (66.7)	
Number of smart devices or wearables owned, n (%)[a]	47	531	578	.049
1	39 (83)	368 (69.3)	407 (70.4)	
≥2	8 (17)	163 (30.7)	171 (29.6)	
Duration of use of smart devices/wearables, n (%) [a]	45	530	575	.19
<1 year	19 (42)	297 (56.0)	316 (55.0)	
1-2 years	18 (40)	153 (28.9)	171 (29.7)	
>2 years	8 (18)	80 (15.1)	88 (15.3)	
Use of smart devices/wearables in partnership with health care provider, n (%)[a]	46	533	579	.17
Yes	3 (7)	73 (13.7)	76 (13.1)	
No	43 (94)	460 (86.3)	503 (86.9)	
Satisfaction with smart devices/wearables, mean (min-max)[b]	4.08 (2-5)	4.07 (1-5)	4.08 (1-5)	.98
Ease of use, mean (min-max)[b]	4.20 (2-5)	4.21 (1-5)	4.21 (1-5)	.92
Expectation confirmation, mean (min-max)[b]	3.78 (1.67-5)	3.89 (1-5)	3.88 (1-5)	.31
Perceived usefulness, mean (min-max)[b]	3.66 (1.50-5)	3.82 (1-5)	3.80 (1-5)	.15
Intention to continue using smart devices/wearables, mean (min-max)[b]	4.22 (1-5)	4.26 (1-5)	4.25 (1-5)	.75

[a]The total values in the rows indicate the number of respondents for that category, which may be lower than the total number of older adults and the general adult population.

[b]The means represent the average of 4 questions that constitute each scale (satisfaction with mobile apps, ease of use, expectation confirmation, perceived usefulness, and intention to continue using mobile apps). Continuous variables were measured on a 5-point Likert scale.

When asked about the types of devices used, the answers also varied between the 2 groups. The most commonly reported devices were bracelets/wristbands. The adults in both the groups who reported using smart devices/wearables did not differ significantly in relation to the duration of use of these technologies and the extent of use in partnership with a health care provider, which was relatively low among the respondents, that is, 7% of the older adults (3/46) and 13.7% (73/530) of the general adult population. As in the case of mobile app use for health, no significant differences were observed in the factors that affect the use of smart devices/wearables (ie, perceived ease of use, perceived usefulness, and expectation confirmation) between the older adult and the general adult population. The overall satisfaction with and the intention to continue using these smart devices/wearables were high. The older adults who used wearables and smart devices reported being very satisfied (mean 4.1 on the 5-point Likert scale), perceived their devices to be user-friendly (mean 4.2), and had a firm intention to continue using them in the future (mean 4.2). Importantly, respondents perceived these devices as relatively useful. About 6 out of 10 users said that they maintained or improved their health status by using digital self-tracking devices. Approximately 66% (31/47) of the older adult users of smart devices/wearables reported they were more informed or more knowledgeable about their health condition due to the use of these devices. For their part, 53% (25/47) of the older adult users said they felt more confident taking care of their health or more autonomous in the management of their condition. Interestingly, feeling less anxious about one's own health and having more informed discussions with a doctor were not perceived as major benefits among the older adult group.

Perception of Smart Devices and Self-tracking Behaviors Among Older Adults

Cronbach alpha was used to assess the reliability of the measures related to satisfaction and use of mHealth technologies, which were included in this study. The results (see Table 5) show that all the measures exceed the .70 threshold of statistical

significance [47]. The validity of the variables was also supported; the square root of the variance shared by each variable and its respective items (diagonal) was greater than the intercorrelations between the variables.

Table 5. Variance shared by the variables considered in this study.[a]

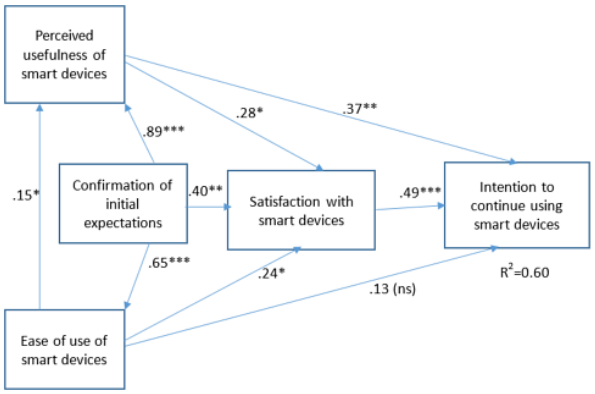
Variables	Number of items	Cronbach α	Perceived usefulness	Perceived ease of use	Confirmation of expectations	User satisfaction	Intention to continue usage
Perceived usefulness	4	.86	*0.82*	0.42[b]	0.79[b]	0.70[b]	0.71[b]
Perceived ease of use	4	.88		*0.84*	0.65[b]	0.62[b]	0.45[b]
Confirmation of expectations	3	.70			*0.83*	0.78[b]	0.63[b]
User satisfaction	3	.88				*0.89*	0.74[b]
Intention to continue usage	3	.93					*0.90*

[a]The diagonal (italicized values in the table) refers to the square root of the variance shared by each variable and its respective items. The values off the diagonal refer to the intercorrelations between the variables. The values in the lower part of the table are a mirror of those in the upper part above the diagonal.

[b]The correlation was significant at $P<.01$.

Partial least squares regression analyses that were performed to test the associations between satisfaction, initial expectations, and intention to continue using smart devices/wearables (Figure 1) showed that all relationships but one were supported, and the model explained 60% of the variance in the dependent variable. These results indicate that expectation confirmation is strongly related to ease of use, perceived usefulness, and user satisfaction, which in turn affect the older adults' intentions to continue using these mHealth technologies.

Figure 1. Results of the partial least squares regression analyses that were performed to test the associations between satisfaction, initial expectations, and intention to continue using smart devices/wearables. * $P<.05$; ** $P<.01$; *** $P<.005$; ns: nonsignificant.



Last, a multinomial logistic regression, including sociodemographic and health status variables, was performed to calculate the odds ratios describing the odds of tracking one's own health using traditional or digital devices compared with the odds of nontracking (reference category) among the older adult group. The traditional .05 criterion of statistical significance was employed for all tests. Addition of the predictors to a model that contained only the intercept significantly improved the fit between model and data; $\chi^2_{34}(n=682)=49.46$, Nagelkerke $R^2=0.11$, $P<.01$.

As indicated in Table 6, our analyses showed no statistically significant differences between the groups (traditional trackers, digital trackers, and nontrackers) in terms of gender, education level, occupation, and perceived health condition. However, significant differences were observed in terms of region ($P=.005$ and $P=.03$ for traditional trackers and digital trackers, respectively, in Alberta) and chronic conditions ($P<.001$ and $P=.003$ for traditional trackers and digital trackers, respectively). Older adults living in the province of Alberta were 4.9 times more likely to be in the digital self-tracking group than the older adults living in other Canadian regions. Compared with older adults living with no chronic condition, older adults with chronic conditions were 0.4 times less likely to be digital self-trackers.

Table 6. Multinomial logistic regression results[a].

Characteristics	Traditional trackers, n=307		Digital trackers, n=121	
	OR[b] (95% CI)	*P* value	OR[b] (95% CI)	*P* value
Intercept	N/A[c]	<.001	N/A	<.001
Gender				
Female	1.176 (0.761-1.817)	.47	1.144 (0.656-1.996)	.63
Region				
Atlantic provinces	1.342 (0.582-3.094)	.49	1.098 (0.379-3.183)	.86
Quebec	1.896 (0.983-3.658)	.06	1.322 (0.585-2.988)	.50
Ontario	1.503 (0.828-2.729)	.18	1.001 (0.478-2.096)	>.99
Prairies	1.270 (0.487-3.309)	.62	0.793 (0.220-2.861)	.72
Alberta	6.053 (1.719-21.312)	.005	4.914 (1.221-19.775)	.03
Education				
Primary and secondary school	1.064 (0.539-2.102)	.66	0.623 (0.274-1.616)	.37
College/CEGEP	1.371 (0.686-2.738)	.37	0.623 (0.541-2.960)	.59
University undergraduate	1.270 (0.675-2.392)	.46	0.832 (0.571-2.675)	.59
Occupation				
Full-time employment	1.078 (0.181-6.422)	.93	0.633 (0.099-4.438)	.67
Part-time employment	1.972 (0.342-11.375)	.45	0.799 (0.118-5.404)	.82
Retired	1.596 (0.333-7.649)	.56	0.699 (0.139-3.525)	.66
Perceived health condition				
Very poor/ poor	0.524 (0.240-1.14)	.10	0.734 (0.275-1.955)	.54
Fair or good	1.094 (0.692-1.730	.70	1.249 (0.697-2.240)	.45
≥1 chronic disease(s)	0.439 (0.281-0.686)	<.001	0.431 (0.245-0.758)	.003

[a]Reference category: nontrackers (n=254).

[b]OR: odds ratio.

[c]N/A: not applicable.

Discussion

Study Relevance

This study investigates older adults' use of mHealth technologies in comparison to that of the general adult population and assesses the pattern of use of these technologies for self-tracking purposes. The surveyed older adult population differed significantly from the general population in relation to the sociodemographic variables. This stresses the importance of having a closer examination of older adults' use of mHealth technologies for self-tracking purposes separately from the general adult population, which would inform future research, practice, and policy efforts in this area.

Principal Findings

Although there were significant differences between older adults and the general adult population in the use of internet and smart technology, a considerable number of older adults reported using them. Specifically, 47.3% (323/682) and 49.8% (340/682) of the older population (65 years and older) reported using a smartphone or a digital tablet, respectively, and 67.8% (314/463) indicated accessing the internet on a daily basis. A large number of Canadian older adults in the community have already acquired these technologies, which presents an opportunity to leverage them beyond basic communication use to support their well-being by enhancing social connectedness and improving the management of their health conditions [48,49].

Despite the comparable good-to-excellent health status reported by older adults and the general adult population, the prevalence of chronic conditions was significantly higher among the older adult group compared to that among the general adult population (*P*<.001), which necessitates close monitoring and management of their health and conditions. Therefore, it is important to leverage existing technologies that can support their health and well-being needs in the community and potentially connect them with caregivers and health care providers. This is particularly relevant in relation to wearables (eg, wristbands, pedometers) and mobile apps that allow users to store and monitor health-related data. Prior research has discussed the important role of technology to support the ability of older adults to remain at home, improve their quality of life and health outcomes, and enhance family caregivers' and health care

professionals' access to relevant information [16,50]. This is in line with the findings of this study that showed a high satisfaction rate with mHealth technologies and favorable conditions for their use.

Nevertheless, this study demonstrates that the potential of mHealth technologies for self-tracking purposes has not been fully captured yet in the context of older adults. Although 62.8% (428/682) of the older adults reported tracking their health measures, the majority did so manually, which may compromise the process, given the risk of losing information and the difficulty in sharing it with health care providers and caregivers. This considerable number of older adults tracking their health measures is indicative of the need and interest among this group to monitor their health. In the absence of the widespread use of personal health records, older adults do not have options for tracking and monitoring their health status but through their own initiative; as such, many seem to resort to the traditional manual recording of their health. This may be an indication of limited knowledge that they may have on mHealth technologies and how they work or a lack of funding and incentives to acquire and use these technologies (ie, from the health care providers, government, and caregivers). Surprisingly though, older adults with one or more chronic conditions appeared to be mostly nontrackers, which raises concerns as to the extent to which mHealth technologies are indeed benefitting the older adults most in need of them.

In light of these findings, it is critical to develop strategies to enhance older adults' awareness and knowledge of the existing mHealth technologies available at their disposal and how to use them and encourage family physicians and allied health professionals to communicate about these options with them. In addition, it is equally important to understand older adults' priorities and self-tracking needs in order to offer technologies suitable to address these needs [32]. This is particularly relevant in light of recent studies in other countries showing that older adults' acceptance of mobile apps can be improved by informing them about the potential benefits of these technologies [51] and that older adults agree to share collected data through in-home monitoring and sensors with professional caregivers and demand participation in decisions about technology [52].

Interestingly, the majority of the mobile apps downloaded by the surveyed older adults consisted of apps used for health and well-being, reflecting a "targeted" use of these technologies by older adults. Around half of the older adults who reported mobile app use in the past 3 months specified using 2 or more of these apps. This is indicative of the perceived benefits of these technologies by older adults and can also reveal a level of comfort and interest in the use of these mobile apps over time. Once older adults start using mobile apps for health, their interest and willingness to use more than one mobile app over a long period of time was confirmed (Table 3 shows higher proportion of older adults reporting >1-year use compared to the general population). Future studies should investigate the motivating factors that facilitate their embracement of mHealth technologies to develop strategies that would enable a broader range of older adults to benefit from them.

It is also important to note that a low proportion of respondents among the older adults and in the general population, that is, 39% (30/77) and 35.2% (436/1238), respectively, indicated either sharing data from mobile apps or using smart devices/wearables in partnership with health care providers. Hence, it appears to be a disconnect between the actual needs and willingness of the older adults in the community to use mHealth technologies and the ability and readiness of health care providers to leverage these tools to support the care provided for these individuals. Despite previous efforts to explore the factors that affect health information technology adoption by older adults in the community [26], we have very limited information about the facilitators and barriers that play a role in bridging this disconnect and enabling more optimal use of mHealth technologies for older adults' care.

The partial least squares regression analyses confirmed that expectation confirmation is strongly related to ease of use, perceived usefulness, and user satisfaction. Hence, it is critical to adequately manage older adults' initial expectations to ensure greater adherence and continued usage of wearables and smart devices. These initial expectations may be considered as the anchor for the subsequent behavior of older adults, and their acceptance and the use of these technologies, and which may be shaped by the environment in which they live. Caregivers and family members, peers, as well as health care providers can play a significant role in shaping these initial expectations and the subsequent benefits that older adults may reap out of using these technologies. Interestingly, the results of this study show that older adults living in Alberta were 4.9 times more likely to be in the digital self-tracking group compared to older adults in other regions. Alberta is a province known to attract young families and is known for its highest rate of workforce growth. This may have implications for older adults living in this province who are surrounded by a younger population heavily immersed in technology and who may have expectations in relation to the role of mHealth technologies in the care for their older persons.

A culture shift in the provision of care to Canadian older adults living in the community is deemed necessary in order to keep up with the development of mHealth technologies and the changing demographics and expectations of patients and their caregivers. This is particularly important in light of the results in this study that show that older adults living with chronic conditions are 0.4 times less likely to be digital self-trackers. This is a "missed opportunity" at the community level as the individuals who may benefit most from mHealth technologies (ie, older adults with chronic conditions) do not seem to be actually using them. Given this state, how can we make this leap and paradigm shift? Evidently, this shift cannot come along without paralleled changes at the health system level in relation to existing policies, reimbursement modalities, and the structure of health care services delivery. In order to optimize the use of mHealth technologies to support older persons in the community, who need and are capable of using them, it is important that health care providers integrate data gathered through these smart devices in the delivery of care to them.

With the recent COVID-19 crisis, we have seen a rapid uptake of virtual care worldwide, which has been catalyzed by a dire

need to provide "remote care" to a vulnerable population (ie, older adults) and a facilitated reimbursement approach. For example, as of May 1, 2020, the Ministry of Health and Long-Term Care in the province of Ontario, one of the largest health jurisdictions in Canada, implemented new temporary fee schedule codes that cover virtual assessments and provision of services [53]. Despite this agile adaptation during the time of crisis, it is equally important to develop long-term plans to leverage technologies to support the care for Canadian older adults, which may require reforms at the health system level.

Before the COVID-19 crisis, we had started to witness unconventional changes in this area in some Canadian provinces with initiatives that allowed patients to leverage wearables and smart devices to support their health. Alberta, for example, had released a personal health record initiative allowing patients to collect and store their own health data by using wearables and smart medical devices and manage authorizations for accessing these data. Other provinces, including Quebec, Nova Scotia, and Saskatchewan, are following this lead with health information portals giving patients more access and control over their health data. These initiatives are promising; however, they have to be paralleled and supported by changes at the policy and reimbursement levels to close the loop and encourage health care providers to endorse new technologies as integrated components in the delivery of health services for older adults and enablers for improved quality of care.

It is worth noting that the consistent high satisfaction of older adults with mobile apps and smart devices/wearables and their intention to continue using them is a positive indication of the evolving expectations of the older adult population and a potential game changer for the future of care for older adults. The results of this study confirm that once mobile apps and smart devices/wearables are used, the perceived ease of use and usefulness of these technologies do not vary by age of the users. As the older adult population continues to grow to include people currently still in the workforce and using technology in their daily lives (eg, mobile apps, smart devices), the demand for more connectedness with health care providers and better response from the health care system in a networked society will likely increase.

Limitations and Future Research

This study presents contributions to an underresearched area on older adults and mHealth technology use. These findings are the first step toward understanding the behaviors and attitudes of older adults toward these technologies. By unveiling the actual prevalence of mHealth technology use among the Canadian older adult population and exploring their familiarity and satisfaction with these technologies, we set the stage for future research to investigate the optimal environment and predictors for their effective use [54].

At present, there are still significant differences between older adults and the general adult population in relation to the use of mHealth technologies. This necessitates a particular focus on older adults in future studies in order to better understand the needs and perceived facilitators and barriers for the use of these technologies this group. However, interestingly, both groups considered in this study demonstrated similarities in terms of limited current use of mobile apps and wearable devices for sharing data in partnership with health care providers. This calls for future research, which extends to the whole population, to better understand the underlying reasons and challenges in this area and study the feasibility and readiness of health care providers to leverage these tools to support the care that they provided to their patients.

Last, it is important to note some limitations associated with the study design and breadth of data. The data set used in this study is from a single country, thereby limiting the generalizability of the findings. In addition, the web-based survey was completed by respondents who had access to the internet, which may preclude representativeness of potential respondents with no internet access. Given the cross-sectional nature of the survey, a full assessment of the predictors of older adults' use of mHealth technologies as well as an evaluation of the variation in their behaviors over time, especially in relation to changes in their health conditions, was not feasible. Furthermore, given the exploratory nature of the study and the focus on mobile apps and smart devices/wearables, limited data were collected on the functional ability of the older adults, their level of independence and health condition, and other sociodemographic characteristics that may play a role in shaping their use of these technologies. Future studies should take these factors into account to better understand the variation in the use of mHealth technologies by older adults in the community and determine the optimal conditions in which these technologies can best benefit them.

Conclusion

The burden of population aging and the associated chronic conditions is observed worldwide. Mobile technologies present an opportunity to address the challenges faced by older adults in relation to their health and the care that they receive. This study shows that a considerable number of older adults are familiar with and use these technologies. Importantly, older adults who use mHealth technologies are highly satisfied with them and plan to continue using them in the future. Understanding why older adults who are familiar with mHealth technologies are not using them would inform progress in this area. In particular, leveraging these mHealth technologies for older adults who need and may benefit from them, in partnership with family physicians and allied health care professionals remains very limited at present. The current development and deployment of various personal health record initiatives in Canada appear as a promising avenue to facilitate bidirectional health information exchanges between health care providers and patients, including older adults.

Conflicts of Interest

None declared.

References

1. Bloom DE. 7 billion and counting. Science 2011 Jul 29;333(6042):562-569. [doi: 10.1126/science.1209290] [Medline: 21798935]
2. Osborn R, Doty MM, Moulds D, Sarnak DO, Shah A. Older Americans Were Sicker And Faced More Financial Barriers To Health Care Than Counterparts In Other Countries. Health Aff (Millwood) 2017 Dec;36(12):2123-2132. [doi: 10.1377/hlthaff.2017.1048] [Medline: 29140737]
3. United N. World population ageing 2017. 2017. URL: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Report.pdf [accessed 2020-10-25] [WebCite Cache ID https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Report.pdf]
4. GBD 2016 Disease and Injury Incidence and 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]
5. The Commonwealth Fund. 2017 Commonwealth Fund International Health Policy Survey of Older Adults in 11 Countries. URL: https://www.commonwealthfund.org/publications/infographic/2017/dec/2017-commonwealth-fund-international-health-policy-survey-older [accessed 2020-10-25]
6. Goethals L, Barth N, Guyot J, Hupin D, Celarier T, Bongue B. Impact of Home Quarantine on Physical Activity Among Older Adults Living at Home During the COVID-19 Pandemic: Qualitative Interview Study. JMIR Aging 2020 May 07;3(1):e19007 [FREE Full text] [doi: 10.2196/19007] [Medline: 32356777]
7. Goodman-Casanova JM, Dura-Perez E, Guzman-Parra J, Cuesta-Vargas A, Mayoral-Cleries F. Telehealth Home Support During COVID-19 Confinement for Community-Dwelling Older Adults With Mild Cognitive Impairment or Mild Dementia: Survey Study. J Med Internet Res 2020 May 22;22(5):e19434 [FREE Full text] [doi: 10.2196/19434] [Medline: 32401215]
8. Li HO, Huynh D. Long-term social distancing during COVID-19: A social isolation crisis among seniors? CMAJ 2020 May 25;192(21):E588. [doi: 10.1503/cmaj.75428] [Medline: 32575063]
9. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
10. Palmer K, Monaco A, Kivipelto M, Onder G, Maggi S, Michel J, et al. The potential long-term impact of the COVID-19 outbreak on patients with non-communicable diseases in Europe: consequences for healthy ageing. Aging Clin Exp Res 2020 Jul;32(7):1189-1194 [FREE Full text] [doi: 10.1007/s40520-020-01601-4] [Medline: 32458356]
11. Canadian Institute for Health Information. How Canada Compares: Results From The Commonwealth Fund's 2016 International Health Policy Survey of Adults in 11 Countries. 2017. URL: https://www.cihi.ca/sites/default/files/document/text-alternative-version-2016-cmwf-en-web.pdf
12. Health Care in Canada, 2011-a Focus on Seniors and Aging. Canadian Institute for Health Information. 2011. URL: https://secure.cihi.ca/free_products/HCIC_2011_seniors_report_en.pdf
13. Statistics Canada. Canada's population estimates: Age and sex, July 1, 2018. 2019. URL: https://www150.statcan.gc.ca/n1/en/daily-quotidien/190125/dq190125a-eng.pdf
14. Canadian Institute for Health Information. Infographic: Canada's Seniors Population Outlook: Uncharted Territory. : CIHI; 2017. URL: https://www.cihi.ca/en/infographic-canadas-seniors-population-outlook-uncharted-territory
15. World Health Organization. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. 2011. URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2020-09-25]
16. Czaja SJ. Long-term care services and support systems for older adults: The role of technology. Am Psychol 2016;71(4):294-301. [doi: 10.1037/a0040258] [Medline: 27159436]
17. Heinz M, Martin P, Margrett JA, Yearns M, Franke W, Yang H, et al. Perceptions of technology among older adults. J Gerontol Nurs 2013 Jan;39(1):42-51. [doi: 10.3928/00989134-20121204-04] [Medline: 23244061]
18. McMurtrey M, Zeltmann S, Downey J, McGaughey R. Seniors and technology: results from a field study. J Comput Inf Syst 2011;51(4):22-30. [doi: 10.1080/08874417.2011.11645498]
19. Jaana M, Paré G. Home telemonitoring of patients with diabetes: a systematic assessment of observed effects. J Eval Clin Pract 2007 Apr;13(2):242-253. [doi: 10.1111/j.1365-2753.2006.00686.x] [Medline: 17378871]
20. Jaana M, Pare G, Sicotte C. Hypertension Home Telemonitoring. 2007;15(1):19-31. [doi: 10.2165/00115677-200715010-00004]
21. Jaana M, Paré G, Sicotte C. Home telemonitoring for respiratory conditions: a systematic review. Am J Manag Care 2009 May;15(5):313-320 [FREE Full text] [Medline: 19435399]
22. Jaana M, Sherrard H, Paré G. A prospective evaluation of telemonitoring use by seniors with chronic heart failure: Adoption, self-care, and empowerment. Health Informatics J 2019 Dec;25(4):1800-1814. [doi: 10.1177/1460458218799458] [Medline: 30247080]
23. Jaana M, Sherrard H. Rural-Urban Comparison of Telehome Monitoring for Patients with Chronic Heart Failure. Telemed J E Health 2019 Feb;25(2):101-108. [doi: 10.1089/tmj.2017.0303] [Medline: 29847242]

24. Kitsiou S, Paré G, Jaana M, Gerber B. Effectiveness of mHealth interventions for patients with diabetes: An overview of systematic reviews. PLoS One 2017;12(3):e0173160 [FREE Full text] [doi: 10.1371/journal.pone.0173160] [Medline: 28249025]
25. Paré G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. J Am Med Inform Assoc 2007;14(3):269-277 [FREE Full text] [doi: 10.1197/jamia.M2270] [Medline: 17329725]
26. Kavandi H, Jaana M. Factors that affect health information technology adoption by seniors: A systematic review. Health Soc Care Community 2020 May 07. [doi: 10.1111/hsc.13011] [Medline: 32378769]
27. Fox S. The Self-Tracking Data Explosion. Pew Research Center. 2013 Jun 03. URL: https://www.pewinternet.org/2013/06/04/the-self-tracking-data-explosion/ [accessed 2020-09-25]
28. Lupton D. Self-tracking, health and medicine. Health Sociology Review 2016 Nov 21;26(1):1-5. [doi: 10.1080/14461242.2016.1228149]
29. McMahon SK, Lewis B, Oakes M, Guan W, Wyman JF, Rothman AJ. Older Adults' Experiences Using a Commercially Available Monitor to Self-Track Their Physical Activity. JMIR Mhealth Uhealth 2016 Apr 13;4(2):e35 [FREE Full text] [doi: 10.2196/mhealth.5120] [Medline: 27076486]
30. 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;3:2055207617740088. [doi: 10.1177/2055207617740088] [Medline: 29942617]
31. Ringeval M, Wagner G, Denford J, Paré G, Kitsiou S. Fitbit-Based Interventions for Healthy Lifestyle Outcomes: Systematic Review and Meta-Analysis. J Med Internet Res 2020 Oct 12;22(10):e23954 [FREE Full text] [doi: 10.2196/23954] [Medline: 33044175]
32. Caldeira C, Chen Y. Seniors and Self-tracking Technology. In: Sayago S, editor. Perspectives on Human-Computer Interaction Research with Older People. New York: Springer Publishing Company; 2019:67-79.
33. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf M, Wei E. "You Get Reminded You're a Sick Person": Personal Data Tracking and Patients With Multiple Chronic Conditions. J Med Internet Res 2015 Aug 19;17(8):e202 [FREE Full text] [doi: 10.2196/jmir.4209] [Medline: 26290186]
34. Sacco G, Lléonart S, Simon R, Noublanche F, Annweiler C, TOVID Study Group. Communication Technology Preferences of Hospitalized and Institutionalized Frail Older Adults During COVID-19 Confinement: Cross-Sectional Survey Study. JMIR Mhealth Uhealth 2020 Sep 18;8(9):e21845 [FREE Full text] [doi: 10.2196/21845] [Medline: 32896832]
35. Siette J, Wuthrich V, Low L. Social Preparedness in Response to Spatial Distancing Measures for Aged Care During COVID-19. J Am Med Dir Assoc 2020 Jul;21(7):985-986 [FREE Full text] [doi: 10.1016/j.jamda.2020.04.015] [Medline: 32674832]
36. Paré G, Leaver C, Bourget C. Diffusion of the Digital Health Self-Tracking Movement in Canada: Results of a National Survey. J Med Internet Res 2018 May 02;20(5):e177 [FREE Full text] [doi: 10.2196/jmir.9388] [Medline: 29720359]
37. Bhattacherjee A. Understanding Information Systems Continuance: An Expectation-Confirmation Model. MIS Quarterly 2001 Sep;25(3):351. [doi: 10.2307/3250921]
38. Hong S, Thong JY, Tam KY. Understanding continued information technology usage behavior: A comparison of three models in the context of mobile internet. Decision Support Systems 2006 Dec;42(3):1819-1834. [doi: 10.1016/j.dss.2006.03.009]
39. Oliver RL. A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. Journal of Marketing Research 1980 Nov;17(4):460 [FREE Full text] [doi: 10.2307/3150499]
40. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
41. 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]
42. Accenture C.. Patients want a heavy dose of digital. 2016. URL: https://www.accenture.com/t20160629T045303__w__/us-en/_acnmedia/PDF-6/Accenture-Patients-Want-A-Heavy-Dose-of-Digital-Infographic.pdf [accessed 2017-10-26]
43. Brown B. 2016 Connected Patient Report. Health Tech Insider. URL: https://healthtechinsider.com/2016/08/25/2016-connected-patient-report/ [accessed 2019-06-14]
44. Growth from Knowledge (GfK). Health and Fitness tracking: Global GfK Survey. 2016 Aug 25. URL: https://cdn2.hubspot.net/hubfs/2405078/cms-pdfs/fileadmin/user_upload/country_one_pager/nl/documents/global-gfk-survey_health-fitness-monitoring_2016.pdf
45. IFOP. Objets connectes et usage des donnees: la perception des Francais. Published online. 2015 Dec 1. URL: https://www.hiscox.fr/sites/france/files/wordpress/2015/09/ifop-rapport-de-synthese-hiscox-objets-connectes.pdf [accessed 2020-11-06]
46. Bowling A. Just one question: If one question works, why ask several? J Epidemiol Community Health 2005 May;59(5):342-345. [doi: 10.1136/jech.2004.021204] [Medline: 15831678]
47. Nunnally J. Psychometric Theory. In: Psychometric Theory. New York: McGraw-Hill; 1978.
48. Fischer SH, David D, Crotty BH, Dierks M, Safran C. Acceptance and use of health information technology by community-dwelling elders. Int J Med Inform 2014 Sep;83(9):624-635 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.06.005] [Medline: 24996581]

49. Morris ME, Adair B, Ozanne E, Kurowski W, Miller KJ, Pearce AJ, et al. Smart technologies to enhance social connectedness in older people who live at home. Australas J Ageing 2014 Sep;33(3):142-152. [doi: 10.1111/ajag.12154] [Medline: 24730370]
50. Khosravi P, Ghapanchi AH. Investigating the effectiveness of technologies applied to assist seniors: A systematic literature review. Int J Med Inform 2016 Jan;85(1):17-26. [doi: 10.1016/j.ijmedinf.2015.05.014] [Medline: 26216463]
51. de Veer AJE, Peeters JM, Brabers AEM, Schellevis FG, Rademakers JJDJM, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv Res 2015 Mar 15;15:103 [FREE Full text] [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
52. Claes V, Devriendt E, Tournoy J, Milisen K. Attitudes and perceptions of adults of 60 years and older towards in-home monitoring of the activities of daily living with contactless sensors: an explorative study. Int J Nurs Stud 2015 Jan;52(1):134-148. [doi: 10.1016/j.ijnurstu.2014.05.010] [Medline: 24951084]
53. Ministry of Health. COVID-19 Temporary Fee Schedule Codes Implemented-Physicians Can Begin to Submit Claims for COVID-19 on May 1, 2020. Claims Services Branch, Ontario Ministry of Health and Long Term Care. 2020 Apr 30. URL: http://www.health.gov.on.ca/en/pro/programs/ohip/bulletins/4000/bul4755.pdf [accessed 2020-10-26]
54. Savoli A, Barki H, Pare G. Examining How Chronically Ill Patients' Reactions to and Effective Use of Information Technology Can Influence How Well They Self-Manage Their Illness. MISQ 2020 Jan 01;44(1):351-389. [doi: 10.25300/misq/2020/15103]

Abbreviations

mHealth: mobile health

Edited by L Buis; submitted 01.10.20; peer-reviewed by H Gewald, B Chaudhry; comments to author 18.10.20; revised version received 23.10.20; accepted 23.10.20; published 27.11.20.

Please cite as:
Jaana M, Paré G
Comparison of Mobile Health Technology Use for Self-Tracking Between Older Adults and the General Adult Population in Canada: Cross-Sectional Survey
JMIR Mhealth Uhealth 2020;8(11):e24718
URL: http://mhealth.jmir.org/2020/11/e24718/
doi:10.2196/24718
PMID:33104517

Original Paper

Once Daily Versus Overnight and Symptom Versus Physiological Monitoring to Detect Exacerbations of Chronic Obstructive Pulmonary Disease: Pilot Randomized Controlled Trial

Ahmed M Al Rajeh[1], BSc, MSc, PhD; Yousef Saad Aldabayan[1], BSc, MSc, PhD; Abdulelah Aldhahir[2,3], BSc, MSc, PhD; Elisha Pickett[4], BSc, MSc; Shumonta Quaderi[2], BSc, MBBS; Jaber S Alqahtani[2,5], BSc, MSc; Swapna Mandal[2,4], BSc, MBBS, PhD; Marc CI Lipman[2,4], BSc, MBBS, PhD; John R Hurst[2,4], BSc, MBBS, PhD

[1]Department of respiratory care, King Faisal University, Al-Ahsa, Saudi Arabia

[2]UCL Respiratory, University College London, London, United Kingdom

[3]Respiratory Care Department, Faculty of Applied Medical Sciences, Jazan University, Jazan, Saudi Arabia

[4]Department of respiratory medicine, Royal Free London NHS Foundation Trust, London, United Kingdom

[5]Department of Respiratory Care, Prince Sultan Military College of Health Sciences, Dammam, Saudi Arabia

Corresponding Author:
John R Hurst, BSc, MBBS, PhD
UCL Respiratory
University College London
Gower Street
London, WC1E 6BT
United Kingdom
Phone: 44 20 7679 2000
Email: j.hurst@ucl.ac.uk

Abstract

Background: Earlier detection of chronic obstructive pulmonary disease (COPD) exacerbations may facilitate more rapid treatment with reduced risk of hospitalization. Changes in pulse oximetry may permit early detection of exacerbations. We hypothesized that overnight pulse oximetry would be superior to once-daily monitoring for the early detection of exacerbations.

Objective: This study aims to evaluate whether measuring changes in heart rate and oxygen saturation overnight is superior to once-daily monitoring of both parameters and to assess symptom changes in facilitating earlier detection of COPD exacerbations.

Methods: A total of 83 patients with COPD were randomized to once-daily or overnight pulse oximetry. Both groups completed the COPD assessment test questionnaire daily. The baseline mean and SD for each pulse oximetry variable were calculated from 14 days of stable monitoring. Changes in exacerbation were expressed as Z scores from this baseline.

Results: The mean age of the patients was 70.6 (SD 8.1) years, 52% (43/83) were female, and the mean FEV1 was 53.0% (SD 18.5%) predicted. Of the 83 patients, 27 experienced an exacerbation. Symptoms were significantly elevated above baseline from 5 days before to 12 days after treatment initiation. Day-to-day variation in pulse oximetry during the stable state was significantly less in the overnight group than in the once-daily group. There were greater relative changes at exacerbation in heart rate than oxygen saturation. An overnight composite score of change in heart rate and oxygen saturation changed significantly from 7 days before initiation of treatment for exacerbation and had a positive predictive value for exacerbation of 91.2%. However, this was not statistically better than examining changes in symptoms alone.

Conclusions: Overnight pulse oximetry permits earlier detection of COPD exacerbations compared with once-daily monitoring. Monitoring physiological variables was not superior to monitoring symptoms, and the latter would be a simpler approach, except where there is a need for objective verification of exacerbations.

Trial Registration: ClinicalTrials.gov NCT03003702; https://clinicaltrials.gov/ct2/show/NCT03003702

(JMIR Mhealth Uhealth 2020;8(11):e17597) doi:10.2196/17597

KEYWORDS

chronic obstructive pulmonary disease; exacerbations; telehealth; CAT; heart rate; oxygen saturation

Introduction

Background

Chronic obstructive pulmonary disease (COPD) is a major global health problem. The World Health Organization estimated the global prevalence in 2016 to be 251 million cases, and COPD is ranked as the fifth leading cause of death [1]. COPD is predicted to become the third most common cause of death by 2030 [2]. Individuals with COPD experience acute deteriorations in respiratory health called *exacerbations* [3], which profoundly affect quality of life and can lead to hospital admission [4]. Prevention and mitigation of exacerbations is a key goal in managing COPD. One method that might improve COPD outcomes is the early identification of exacerbations. Prompt access to therapy at exacerbation onset is associated with faster symptom recovery [5]. Delay in accessing care may be because of the difficulty in differentiating day-to-day symptom variations from exacerbations [6]. Therefore, there has been interest in the utility of monitoring physiological variables to support the earlier detection of exacerbations [7]. Early objective identification of exacerbations would also be of value in research.

The utility of measuring physiological variables to detect COPD exacerbations remains unclear [8]. In our 2010 pilot study, we combined daily monitoring of symptoms, heart rate, and oxygen saturation [9]. A composite score created from these variables changed significantly just before exacerbation onset and could be used to monitor recovery. Measurements of heart rate and oxygen saturation in this study were taken once daily, and one-off readings could be affected by non–exacerbation-related factors such as exercise, medication, and anxiety. We hypothesized that monitoring patients' pulse oximetry overnight and removing the effects of non–exacerbation-related factors would provide earlier exacerbation detection.

Objectives

This study aims to evaluate whether measuring changes in heart rate and oxygen saturation overnight is superior to once-daily monitoring of both parameters and to assess symptom changes in facilitating earlier detection of COPD exacerbations.

Methods

Participants and Recruitment

Participants in COPD clinics and pulmonary rehabilitation classes were approached between September 2016 and January 2018 at 3 different sites in London, United Kingdom. Ethical approval was obtained from the local committee at the Royal Free Hospital and the UK Health Research Authority (16/LO/1120). The study was registered at ClinicalTrials.gov (NCT03003702). All the participants provided written informed consent. The inclusion criteria were a confirmed diagnosis of COPD (smoking history ≥10 pack years and postbronchodilator forced expiratory volume in one second/forced vital capacity $FEV_1/FVC<0.7$), one or more self-reported moderate or severe COPD exacerbations in the previous 12 months, and the ability to use study equipment and attend appointments. The exclusion criteria were an existing diagnosis of obstructive sleep apnea (either self-reported or resulting from STOP-Bang or Epworth questionnaires [10,11]) and significant comorbidity preventing participation (such as poor peripheral perfusion).

Study Procedures

At recruitment, we collected demographic and clinical data and performed postbronchodilator spirometry (FEV_1 and FVC). Subjects were randomized using sealed envelopes; each envelope contained a card indicating once-daily or overnight measurement. The researcher instructed the participants on how to use the pulse oximetry equipment based on their allocation wristband pulse oximeter (Nonin 3150) for overnight monitoring (measurement recorded every 4s), or finger pulse oximeter (Nonin G92) for once-daily (morning) monitoring. Patients were also instructed on the use of a peak expiratory flow (PEF) meter and the Chronic obstructive pulmonary disease assessment test (CAT) questionnaire, and how to record the results on a diary card. Subjects were monitored for 6 months or until they recovered from the first exacerbation (whichever was earlier).

Data recorded in the first week were not included in the analysis. During the subsequent 2 weeks (the *stable state*), participants were closely monitored to ensure they were using the equipment properly and recording data accurately. Participants were then instructed to call the researcher if they developed an exacerbation or if they developed any medical problem resulting in hospitalization. Monitoring was continued through the exacerbation until clinical recovery, at which point the equipment was returned and participation was complete. The equipment was removed, and the study was completed for any participant who did not develop an exacerbation within 6 months. In the absence of an exacerbation, in-person visits were scheduled at 2 weeks, 3 months, and 6 months. An exacerbation was defined as the need for oral corticosteroids or antibiotics, as judged by the patient's clinician or self-management plan.

At the end of the study, a 10-question acceptability survey was administered to participants allocated to the overnight monitoring group to evaluate their acceptance of continuous overnight pulse oximetry monitoring. The questionnaire covered willingness to use the monitoring equipment, effect on sleep quality, and convenience of the device. The highest possible score was 90 (the higher the score, the greater the acceptance).

Pulse Oximeter

Participants allocated to the overnight monitoring (Nonin3150 pulse oximeter) group were instructed to wear the device at night before sleeping and to remove it as soon as they woke up (confirmed by the daily *date and time stamp* from the device and then compared with the date on the diary card).

Participants allocated to the once-daily (morning) monitoring (Nonin G92 pulse oximeter) group were instructed to take their measurements before morning medication, after 10 min of rest, and with the measurement recorded after 10 seconds of using the device.

The accuracies of Nonin G92 and Nonin 3150 were assessed by the manufacturer. The devices comply with ISO 80601-2-61 and have a stated accuracy of 70% to 100% SpO_2 (SD 2%) [12,13].

Power Calculation

The primary outcome was the difference in time to receive treatment from symptoms (CAT score) exceeding baseline compared with pulse oximetry exceeding baseline. A power calculation based on 2 previous data sets [11,14] suggested that 64 patients with COPD were required to capture 44 COPD exacerbations.

Statistical Analysis

Data collected from each participant were reviewed by a respiratory therapist. Participants who had less than 2 weeks of data during the stable state were excluded from the analysis. Participants with 2 or more consecutive days of missing data during the preexacerbation phase (2 weeks before exacerbation) were excluded. Normally distributed data are reported as mean (SD) and nonparametric data as median (IQR). Groups were compared using the *t* test, Mann-Whitney *U* test, or Chi-square test as appropriate, and relationships between variables were assessed using the Pearson or Spearman Rank coefficient. Analysis of variance (ANOVA) and Kruskal-Wallis tests were used to assess the differences between 3 different periods (stable state, preexacerbation treatment, and post exacerbation). The baseline (stable state) for each patient was set by calculating the mean and SD over 2 weeks for each variable (average for each day for 14 days). Data were converted to Z score values by subtracting each measure (average per day [X]) from the corresponding mean baseline value (x) and dividing by the baseline SD as follows:  . Any variable outside the 95% CI of the baseline mean (1.96 SD) was considered statistically significant at $P<.05$.

We calculated sensitivity, specificity, and positive predictive values by examining changes in heart rate, oxygen saturation, and peak flow in a second 2-week period when the patient was stable (not the same stable period as used to calculate the baseline).

The acceptability score is presented as mean (SD). Univariate and multivariate linear regression analyses were applied to identify patient characteristics associated with the acceptability of overnight home pulse oximetry.

Results

Patient Characteristics

The study flowchart is shown in Figure 1. In total, we approached 186 patients with COPD for participation in the study, and 47.3% (88/186) agreed. Thus, 88 patients with COPD were recruited and randomized to either the once daily or continuous overnight monitoring groups. Of the 88 patients, 5 patients from the overnight group were subsequently excluded from the study because they later required oxygen therapy or were diagnosed with obstructive sleep apnea. Overall, 16 and 18 patients, respectively, in the overnight and once-daily groups completed follow-up, with 13 and 14 exacerbations available for analysis, respectively.

The baseline characteristics of the 83 patients who received their allocated intervention are presented in Table 1. The mean age was 70.6 (SD 8.1) years, 52% (43/83) were female, and the mean percentage FEV_1 was 53.0% (SD 18.5%). A majority of the patients were ex-smokers 77% (64/83). The median (IQR) number of COPD exacerbations within the past 12 months was 2.0 (1.0- 3.0). We had a higher-than-anticipated dropout rate of 59%, as discussed below. Table 1 also reports the baseline data from the 27 patients in whom exacerbations were successfully monitored. No statistically significant differences were found between the once-daily and the overnight groups, suggesting that the groups were similar at baseline. The successfully monitored participants were also generally similar to the patients who had dropped out of the study.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. OSA: obstructive sleep apnea; O_2: oxygen.

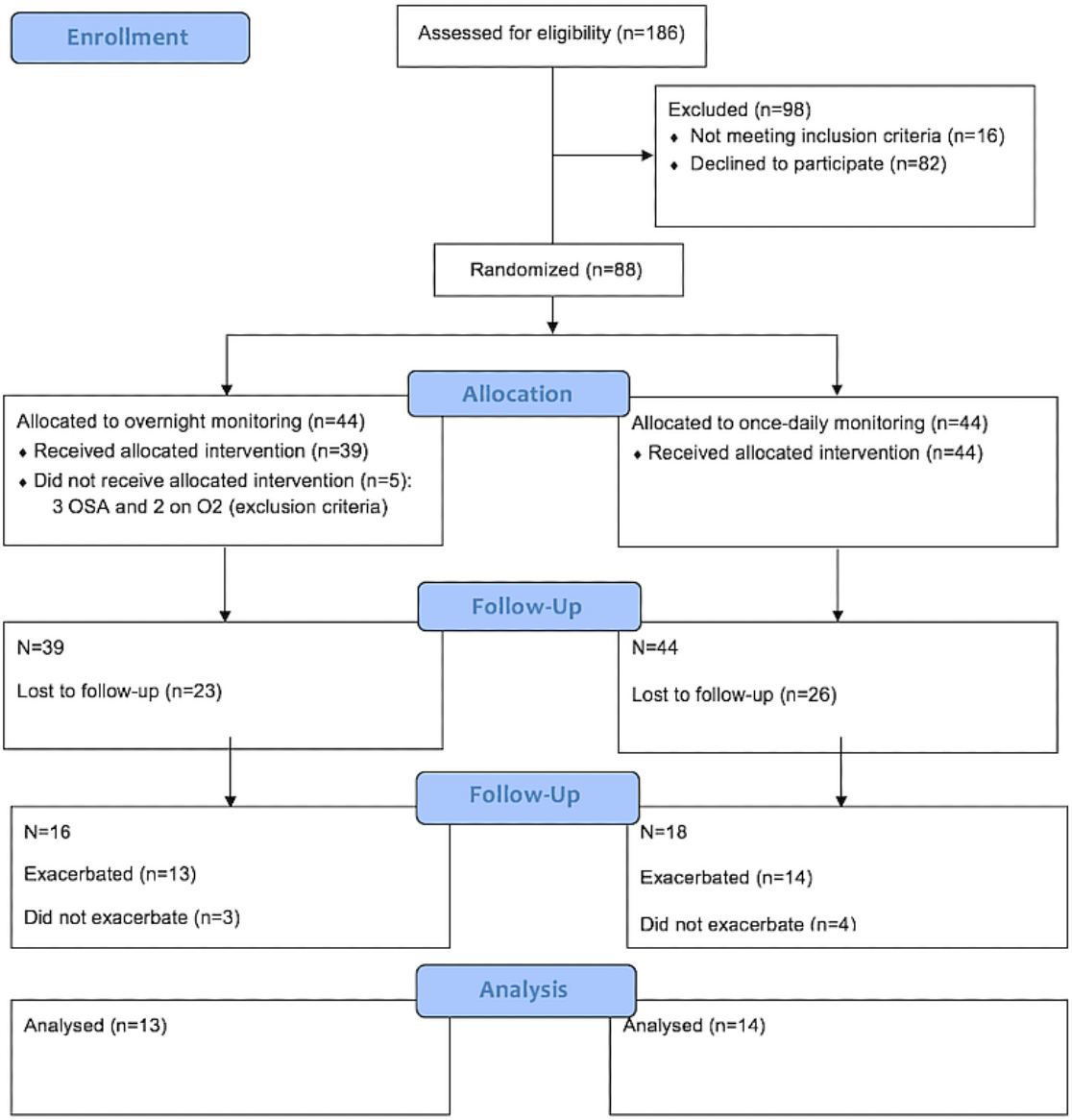


Table 1. Patient demographics and clinical measures.[a]

Characteristics patient demographics and clinical measures	All patients (n=83)	Once-daily arm (n=14)	Overnight arm (n=13)	*P* value
Age (years), mean (SD)	70.6 (8.1)	72.2 (2.6)	70.7 (2.9)	.68
Gender (Female), n %	43 (52)	10 (71)	6 (46)	.18
BMI (kg/m^2), mean (SD)	26.7 (5.9)	25.6 (1.4)	25.5 (1.9)	.98
Smoking status, n (%)				.92
Ex-smoker	64 (77)	11 (79)	10 (77)	
Current smoker	19 (23)	3 (21)	3 (23)	
Do you live with someone? n (%)	43 (52)	7 (50)	8 (62)	.55
FEV^b_1 L, mean (SD)	1.2 (0.4)	1.0 (0.3)	1.1 (0.5)	.63
FEV_1%, mean (SD)	52.9 (18.6)	53.5 (17.7)	44.8 (18.2)	.52
FVC^c L, mean (SD)	2.6 (0.8)	2.3 (0.7)	2.5 (0.9)	.30
FVC %, mean (SD)	83.7 (21.7)	82.9 (17.7)	77.8 (15.6)	.52
FEV_1/FVC %, mean (SD)	51.9 (11.7)	52.6 (14.5)	50.8 (10.7)	.72
MRC Breathlessness Score, mean (SD)	2.9 (0.8)	2.6 (0.2)	2.9 (0.2)	.34
Number of exacerbation/previous year, median (IQR)	2 (1-3)	1.5 (1-2)	2.0 (1-3)	.43
Number of hospitalization/previous year, median (IQR)	0 (0-0)	0 (0-0)	0 (0-0)	.37
Charlson comorbidity index, mean (SD)	4.1 (1.2)	4.0 (0.3)	4.5 (0.2)	.26

[a]Baseline data reported as mean (SD) or median (IQR) unless stated otherwise. *P* value is comparison between once-daily and overnight arms. Data analysis was performed using Statistical Package for Social Sciences (SPSS), Version 24.

[b]FEV1: forced expiratory volume in 1 second.

[c]FVC: forced vital capacity.

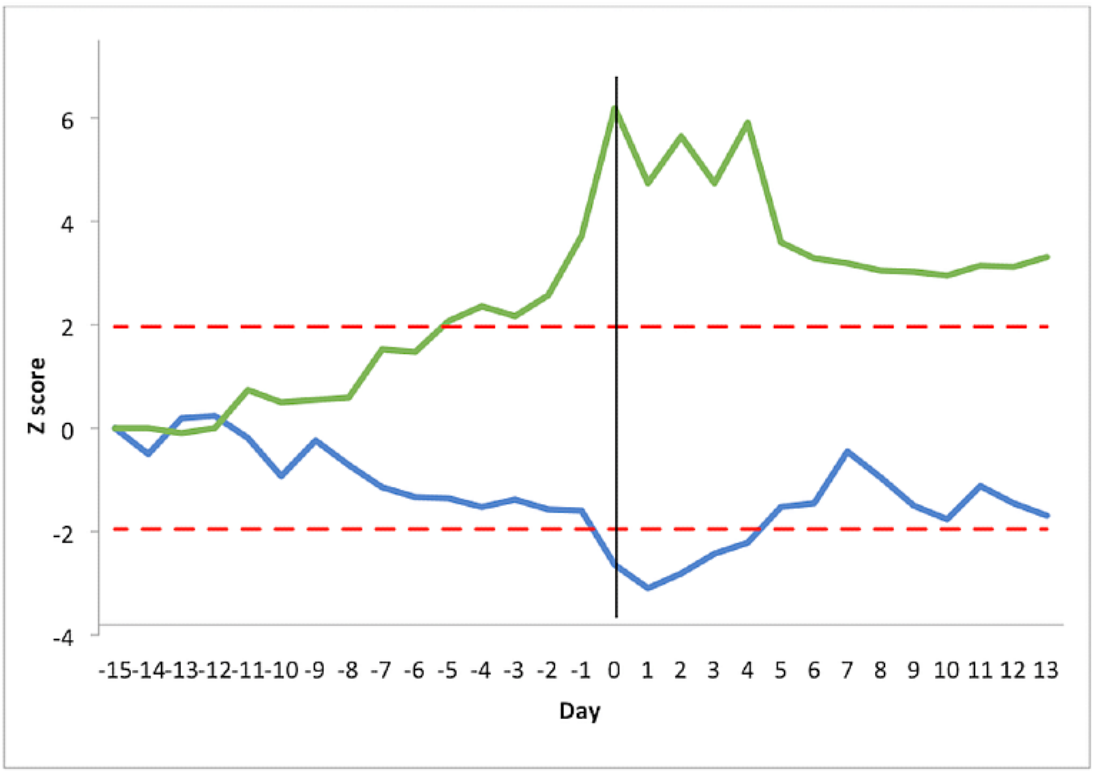
Changes in Symptoms at Exacerbation of COPD

First, we examined changes in the CAT score to understand the symptom changes during exacerbations. Both groups measured the CAT score once each day, and therefore the data from both groups were combined. Figure 2 shows the CAT score from 2 weeks before exacerbation (day −1 to day −14), day 0 (defined as the start of treatment), and for 2 weeks subsequently, as the patient recovered (day 1 to day 13). Day −15 on the graph represents the mean of the stable period (2 weeks baseline monitoring, as described above).

The mean CAT score for the 27 patients with COPD at baseline was 15.6 (SD 0.4) points. Any score greater than or equal to 1.96 SD units away from the mean was considered statistically significant at $P<.05$. Figure 2 demonstrates that the CAT score crossed this statistical threshold from day −5 and did not return to baseline on day 13. The increase in CAT score was greatest on day 0 (the day of treatment initiation), with a maximum change of +6.2 SD (=7 points).

The minimal clinically important difference (MCID) for the CAT questionnaire is 2 points. The mean CAT score for the 27 patients with COPD increased above the MCID from day −5 (=18.3 points) to day 12 (=18.2 points). The difference in mean CAT scores between the 3 phases was statistically significant (stable state=15.6 points, pre exacerbation=17.4 points, post exacerbation=19.7 points, $P<.001$). Thus, patients had statistically and clinically elevated CAT scores from 5 days before receiving treatment for exacerbation to 12 days after treatment initiation. The variability of CAT score increased before exacerbation compared with baseline (SD 2.34 vs 1.46, $P=.004$). From this, we concluded that the exacerbations we detected had been associated with significant changes in symptoms and that it was appropriate to go on and examine changes in PEF and pulse oximetry.

Figure 2. Chronic obstructive pulmonary disease assessment test score (green line) and peak expiratory flow (blue line) changes pre and post exacerbation in 27 patients with chronic obstructive pulmonary disease. The y-axis represents Z score relative to the patient's baseline mean. The x-axis represents the time expressed in days. Day −15 is the mean of the stable period, days −14 to −1 is the preexacerbation period, day 0 is the day of initiation of treatment for exacerbation, and days 1 to 13 is the postexacerbation recovery period. The red lines represent the threshold limits of ±1.96 SD.



Changes in PEF at Exacerbation of COPD

Next, we examined changes in PEF to confirm that changes in symptoms were associated with changes in lung function. The mean PEF for the 27 patients with COPD (both groups) at baseline was 214 (SD 3.7) $Lmin^{-1}$. As illustrated in Figure 2, the PEF decreased gradually before the treatment of exacerbation and fell below 1.96 SD between day 0 and day 4, with the maximal change at day +1.

ANOVA analysis demonstrated a statistically (but not clinically) significant difference in PEF between the 3 different phases (stable state=213.8 $Lmin^{-1}$, pre exacerbation=201.6 $Lmin^{-1}$, post exacerbation=198.0 $Lmin^{-1}$, *P*=.001). Post hoc, the differences were significant between the stable phase vs preexacerbation periods (*P*=.02) and stable vs postexacerbation periods (*P*=.002). There was no difference in PEF variability between the stable and preexacerbation phases (13.81 vs 15.54 $Lmin^{-1}$, *P*=.49).

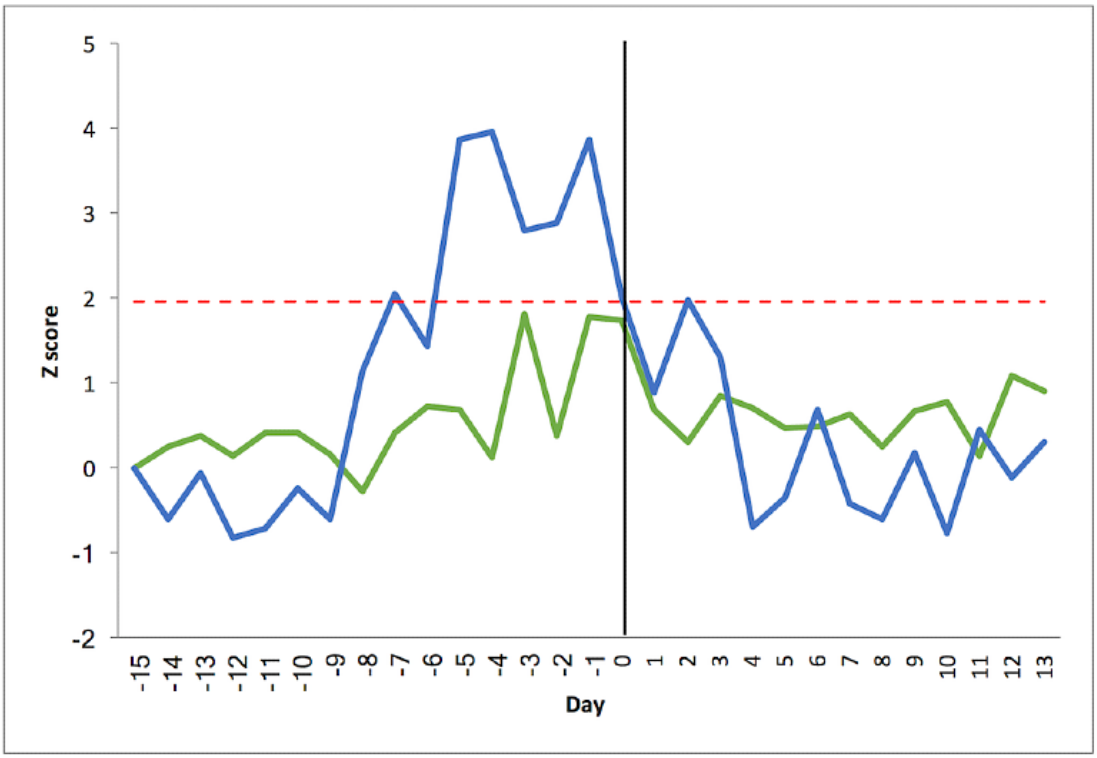
In summary, the exacerbations we captured were associated with typical changes in symptoms and PEF. We next went on to examine changes in the variables monitored differently between the 2 groups: heart rate and oxygen saturation.

Changes in Heart Rate at Exacerbation of COPD

Figure 3 shows the heart rate changes in the once-daily group (n=14) and the overnight group (n=13). The mean stable heart rate for the once-daily group was 77.1 (SD 3.6) beats min^{-1}. At no point did changes in heart rate in the once-daily group cross the threshold of greater or less than 1.96 SD units away from the stable mean. The maximum change occurred at day −1 (SD 1.7), representing an increase of 7 beats min^{-1}. An ANOVA test showed a statistically (but not clinically) significant difference between the 3 periods (stable state=77.1 beats min-1, pre exacerbation=76.7 beats min^{-1}, post exacerbation=81.2 beats min-1, *P*=.007). The difference was significant between post exacerbation vs the stable phase (*P*=.02) and post exacerbation vs pre exacerbation (*P*=.01). In the once daily group, the variability (SD) in heart rate did not differ between the baseline and preexacerbation phases (4.50 vs 5.50, *P*=.28).

The mean heart rate for the overnight group in the stable period was 70.0 (SD 1.8) beats min^{-1}. When compared with the once-daily group, the stable state SD in the overnight group was significantly smaller (1.8 vs 3.6 min^{-1}*P*<.001). In the overnight group, the heart rate did cross the significance threshold and was consistently above this from day −5 to day 0 with a maximal increase of 10 beats min^{-1} (at day −4, equivalent to 3.95 SD). There was a significant difference between the 3 periods (stable state=70.0 min^{-1}, pre exacerbation= 73.9 min^{-1}, post exacerbation=67.6 min^{-1}, *P*=.001). The difference was significant between the preexacerbation period versus the stable state and preexacerbation vs postexacerbation phase (*P*=.04 and *P*=.001, respectively). Heart rate variability was statistically higher in the preexacerbation phase compared with baseline (7.00 vs 4.01 min^{-1}, *P*=.04).

Figure 3. Heart rate changes pre and post exacerbation for the 27 patients with chronic obstructive pulmonary disease. The y-axis represents the Z score relative to the patient's baseline mean. The x-axis represents the time expressed in days. Day −15 is the mean of the stable period, days −14 to −1 is the preexacerbation period, day 0 is the day of initiation of treatment for exacerbation, and days 1 to 13 is the postexacerbation recovery. The red line represents the threshold limit of +1.96 SD. The once-daily group is represented by the green line and the overnight group by the blue line.

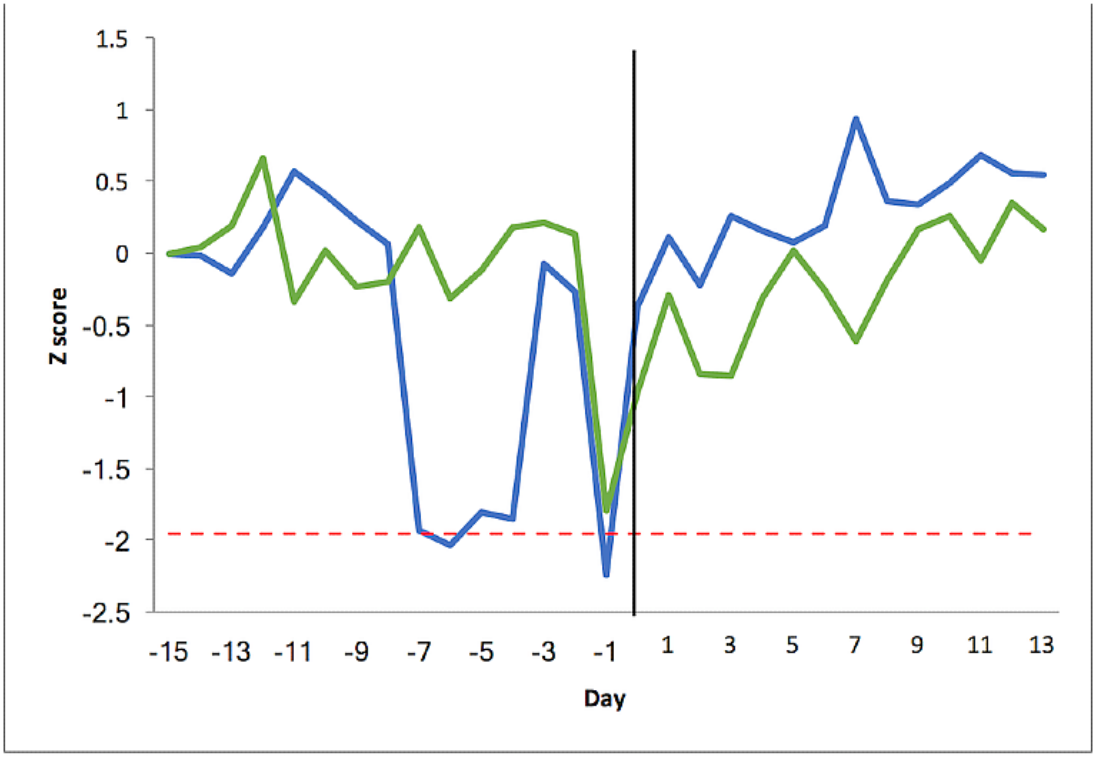


Changes in Oxygen Saturation at Exacerbation of COPD

We next examined trends in oxygen saturation through the time course of an exacerbation. Figure 4 shows the SpO_2% variability in the once-daily measurement group (n=14) and the overnight measurement group (n=13). The mean oxygen saturation during the stable period in the once-daily group was 94.0% (SD 0.81%). SpO_2% variability remained within the stable range in the once-daily group with a maximum reduction of −1.80 SD (−2%) at day −1. There was no significant difference between the 3 periods (stable state=94.0%, pre exacerbation=93.9%, post exacerbation=93.2%, *P*=.09). Variability in the oxygen saturation signal between the stable and preexacerbation phases was not statistically significant (1.12 vs 1.53%, *P*=.13).

The mean oxygen saturation in the stable phase in the overnight group was 91.0% (SD 0.36%). When compared with the once-daily group, the stable state SD in the overnight group was smaller (0.36 vs 0.81% *P*=.002). The SpO_2% decreased more than 1.96 SD in the overnight group on 3 days: −7, −6, and −1 (−1.93 SD, −2.04 SD, and −2.25 SD, respectively), equivalent to a maximal reduction of 1.2%, but these changes did not occur over a consistent period. The means of the 3 phases (stable, 91.0%; pre exacerbation, 91.0%; and post exacerbation, 91.3%) were not statistically significant (*P*=.26) or clinically different. The variability (SD) of the oxygen saturation signal in the overnight group was not significantly different in the preexacerbation phase compared with baseline (1.13% vs 1.01%, *P*=.58).

Figure 4. Oxygen saturation changes pre and post exacerbation for the 27 patients with chronic obstructive pulmonary disease. The y-axis represents the Z score relative to the patient's baseline mean. The x-axis represents the time expressed in days. Day −15 is the mean of the stable period, days −14 to −1 is the preexacerbation period, day 0 is the day of initiation of treatment for exacerbation, and days 1 to 13 is the postexacerbation recovery. The red line represents the threshold limit of +1.96 SD. The once-daily group is represented by the green line and the overnight group by the blue line.



Changes in a Composite Oximetry Score at Exacerbation of COPD

We have previously demonstrated that an exacerbation is typically associated with unidirectional changes in the pulse oximetry signal, specifically an increase in heart rate and a decrease in oxygen saturation, and demonstrated that a composite oximetry score calculated as (Z_{HR}–Z_{SPO_2}) performs better at detecting changes at exacerbation than either variable alone [15]. Therefore, we proceeded to calculate this score for both groups. Figure 5 shows the composite score for the once-daily group (n=14) and the overnight group (n=13). In the once daily group, the composite score increased more than 1.96 SD away from the baseline mean on day −1 and day 0, with a maximum change of 3.58 SD at day −1. In contrast, in the overnight group, the composite score crossed the threshold consistently between days −7 and day 0, with a maximum change of 6.11 SD on day −1.

The composite score crossed the threshold for abnormality in 11 of 13 patients during the 2 weeks before the exacerbation. While examining the same patients in a 2-week period when stable, the score became abnormal at some point in 3 of 11 patients. Thus, the sensitivity, specificity, and positive predictive value of the score in detecting an exacerbation were 84.6%, 81.8%, and 91.7%, respectively. A summary of the results is provided in Table 2.

Figure 5. Composite pulse oximetry score changes pre and post exacerbation for the 27 patients with chronic obstructive pulmonary disease. The y-axis represents the Z score relative to the patient's baseline mean. The x-axis represents the time expressed in days. Day −15 is the mean of the stable period, days −14 to −1 is the preexacerbation period, day 0 is the day of initiation of treatment for exacerbation, and days 1 to 13 is the postexacerbation recovery. The red line represents the threshold limit of +1.96 SD. The once-daily group is represented by the green line and the overnight group by the blue line.

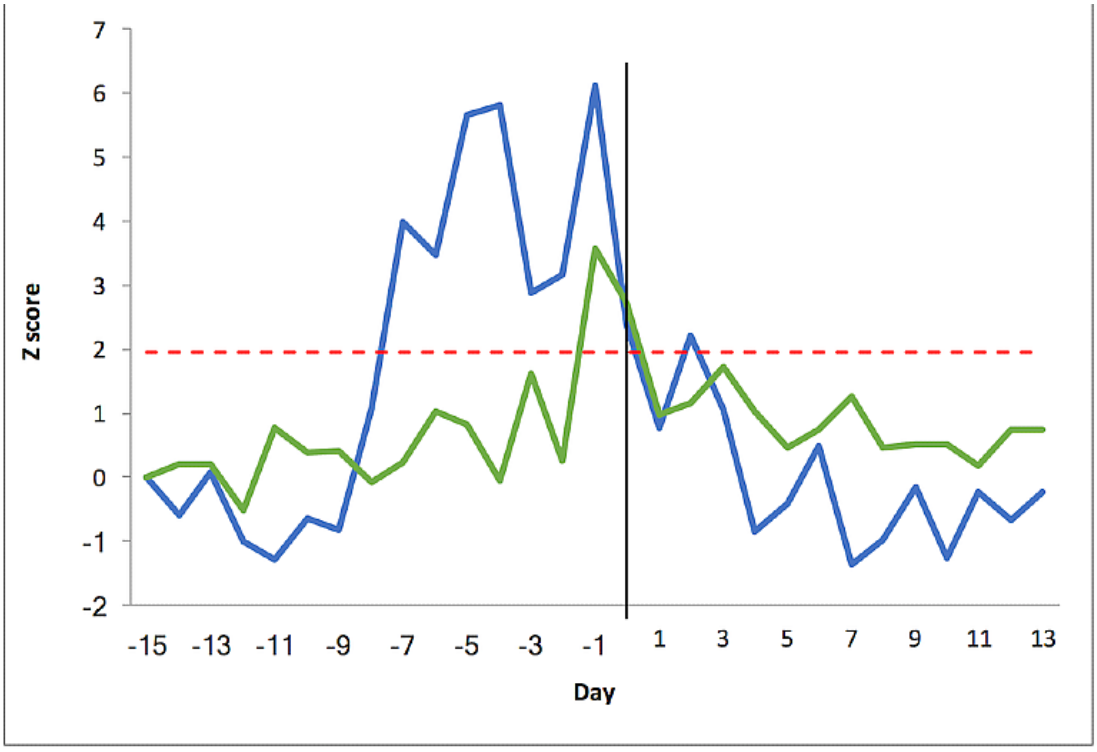


Table 2. Summary table reporting the days when variables showed statistically significant variation from baseline based on a difference of 1.96 SD units or more away from the stable-state mean.

Variable	Combined groups (n=27 patients)	Once-daily group (n=14)	Overnight group (n=13)	*P* value[a]
CAT score	day –5 to day +13	N/A[b]	N/A	<.001
Peak expiratory flow	day 0 to day +4	N/A	N/A	.001
Heart rate	N/A	None[c]	day –7, –5 to day 0, and day +2	.007 (once daily group); .001 (overnight group)
Oxygen saturation	N/A	None	day –7, –6, and –1	.09 (once daily group); .26 (overnight group)
Composite pulse-oximetry score	N/A	day –1, 0	day –7 to day 0, and day +2	.299 (once daily group); .007 (overnight group)

[a]*P* value is the difference in means between the 3 different periods (stable, pre exacerbation, and post exacerbation).

[b]N/A: not applicable.

[c]"None" means that the variable did not cross the threshold of more than 1.96 SD units during the pre- and postexacerbation periods.

Time Difference to Exacerbation Between Symptoms and Physiology

Our prespecified primary analysis was used to assess the time difference to initiate treatment from when a change in symptoms (CAT) was first statistically different from baseline compared with changes in physiological variables (PEF, HR, SpO_2, and composite score). First, we combined both groups to assess the difference between symptoms and PEF (n=27) and found no difference between the groups. The median (IQR) time to receive treatment from symptoms becoming abnormal was 2.0 (0-10) days, whereas for PEF, it was 0.5 (0-5) days (*P*=.15). The results for the differences between symptoms and physiological variables in the once daily and overnight groups are reported in Table 3. None of the pulse oximetry monitoring approaches were superior to the use of CAT alone.

Table 3. Time difference to treatment initiation at exacerbation between symptoms and physiological variables first becoming abnormal.

Groups	CAT[a]–HR[b]	*P* value	CAT—Oxygen Saturation	*P* value	CAT—Composite score	*P* value
Once-daily	1 (0-9) vs 0 (0-0.5) days	.41	1 (0-9) vs 0 (0-2) days	.13	1 (0-9) vs 6.5 (2-11) days	.09
Overnight	5 (0-11) vs 5 (0.5-7.5) days	.78	5 (0-11) vs 0 (0-3.5) days	.13	5 (0-11) vs 6 (2-10) days	.76

[a]CAT: Chronic obstructive pulmonary disease assessment test.

[b]HR: Heart rate.

Finally, we compared the time between change in physiological signal becoming abnormal and receiving treatment between the 2 different monitoring groups. The results are reported in Table 4. There was a statistically significant difference in time between the overnight group and the once-daily group for heart rate (median 5 [0.5-7.5] days vs 0 [0-0.5] days, *P*=.03), but not for the other variables or composite score. This demonstrates that where objective verification of exacerbation is required, overnight monitoring of heart rate has the greatest potential to detect or assist in the detection of exacerbations of COPD and is significantly better than once-daily measurement.

Table 4. Time difference to receive treatment between once-daily and overnight measurment groups.

Groups	CAT[a]	*P* value	Heart rate	*P* value	Oxygen sat	*P* value	PEF[b]	*P* value	Composite score	*P* value
Once-daily	1 (0-9) days	.33	0 (0-0.5) days	.03	0 (0-2) days	.64	0 (0-1.5) days	.12	6 (2-11) days	.79
Overnight	5 (0-11) days		5 (0.5-7.5) days		0 (0-3.5) days		5 (0-8) days		6 (2-10) days	

[a]CAT: Chronic obstructive pulmonary disease assessment test.

[b]PEF: peak expiratory flow.

Patients' Acceptance of Overnight Continuous Pulse Oximetry Monitoring

A total of 29 participants completed the acceptability questionnaire, with a response rate of 97%. Overall, patients had moderate acceptability, with a mean score of 59.8/90 (12.6). Univariate linear regression showed that patients who had a higher Charlson Comorbidity Index score and those who were living alone had higher acceptability (β=.453 *P*=.01, and β=.424 *P*=.02, respectively); thus, acceptability was higher in frailer patients (in whom this technology may potentially be most beneficial).

Discussion

We conducted a randomized trial to compare the ability of overnight vs once daily monitoring of pulse oximetry to provide earlier detection of COPD exacerbations and to assessment of changes in symptoms and PEF. Early detection of exacerbations is important because it could facilitate prompt access to therapy, with faster resolution of exacerbations and decreased risk of hospitalization [5]. Our central hypothesis was that monitoring overnight pulse oximetry would remove non–exacerbation-related effects from the oximetry signal and, therefore, be more effective.

Our key findings are, first, that the time to treatment initiation from changes in symptoms vs changes in pulse oximetry first becoming abnormal were not statistically significant. However, patients waited an average of 5 days before treatment, and therefore, objectively monitoring symptoms or use of overnight oximetry may prompt patients to seek attention. Second, overnight measurement of heart rate may permit earlier detection of exacerbations compared with once-daily measurement because the SD of the stable state mean is smaller, permitting detection of smaller changes (increased signal-to-noise ratio). Third, changes in heart rate were more useful than changes in oxygen saturation or PEF. Finally, a combined oximetry score calculated by subtracting the Z score of change in oxygen saturation from the Z score of change in heart rate had a positive predictive value over 90% for the detection of exacerbation. The findings of this pilot study support our hypothesis that nocturnal pulse oximetry could facilitate earlier objective detection of COPD exacerbations.

Respiratory Symptoms

COPD exacerbations are defined by a change in respiratory symptoms above the day-to-day variation [16]. Seemungal was the first to report that symptoms changed significantly during the seven-day period preceding an exacerbation [6]. Aaron et al [17] went on to describe two patterns of exacerbation symptom onset. Our findings also show that symptoms increased above baseline 5 days before the treatment of exacerbation. In our study, we used the CAT questionnaire to assess symptom severity. A previous study by Lee et al investigated the use of the CAT questionnaire in the prediction of COPD exacerbations [3] and was able to identify patients at risk of developing an exacerbation (area under the curve=0.83). A 2-point change in CAT is clinically significant [18]. Previous studies have shown that the mean CAT score increased above 2 points from baseline at the onset of exacerbation [19]. Our findings are consistent with this, as the mean CAT score in our study increased above baseline by more than 2 points from 5 days before initiation of treatment for exacerbation. However, although the CAT questionnaire is an objective tool for assessment of symptoms, each question is answered based on patients' self-perception, and there remains no objective way to confirm or exclude an exacerbation.

Physiological Variables

In our study, we monitored 3 physiological variables: PEF, heart rate, and oxygen saturation. The evidence for using physiological parameters to identify exacerbations is controversial, reflecting the poor quality of existing studies [7]. The approach used in previous studies was generally once-daily monitoring, and few studies have reported the magnitude of change in physiological variables (which is essential to define alarm limits in clinical practice). We studied two different monitoring approaches and examined differences in day-to-day variation in physiological parameters during 3 different phases (stable, pre exacerbation, and post exacerbation). We show that for heart rate and PEF, but not oxygen saturation, the mean value during the preexacerbation phase was significantly different from that during the stable phase. In addition, variability in the heart rate signal increased as patients first became unwell. These findings support the hypothesis that COPD exacerbations are associated with changes in cardiorespiratory physiology and that monitoring these variables may facilitate the earlier identification of exacerbations.

PEF is a simple tool to assess lung function, and measurements of PEF correlate with FEV_1 [20,21]. The PEF in our study decreased during the preexacerbation phase and was significantly lower than baseline on the day of initiation of treatment for exacerbation. The maximum decrease in PEF during the exacerbation was 34 $Lmin^{-1}$. Seemungal et al reported that PEF decreased below baseline at exacerbation with a median change of 6.6 $Lmin^{-1}$ [22]. In another study conducted by this group, PEF decreased significantly from baseline on the day of exacerbation with a median change of 8.6 $Lmin^{-1}$ [6]. Even though these changes are statistically significant, a decrease in PEF of this magnitude is not clinically significant and is not accurately measurable with current PEF devices. There is no accepted MCID for PEF at exacerbation of COPD.

Heart rate and oxygen saturation in our study were monitored once daily (in the morning) or continuously overnight. Our findings show that the mean heart rate in both groups increased significantly during the preexacerbation phase compared with the stable state. In the overnight group, the heart rate crossed the significance threshold from day −5 for 5 consecutive days. The mean heart rate in the once-daily group increased by 7 min^{-1}, consistent with results from other studies [9,23]. The maximum increase in heart rate in the overnight group was 10 min^{-1}. Importantly, and as we had hypothesized, the variation of heart rate in the overnight group was smaller compared with the variation in the once-daily group (SD 1.8 vs 3.6 min^{-1}), which supports the hypothesis that overnight monitoring of heart rate might enhance the detection of exacerbations by improving the signal-to-noise ratio.

The use of oxygen saturation to support detection and monitoring recovery of community-treated COPD exacerbations has been previously examined. In a pilot study, we reported that oxygen saturation decreased by a mean of 1.2% 2 days into the exacerbation [9]. Other studies have reported that oxygen saturation decreased 1 to 3 days before exacerbation onset by 1% to 2% [23-25]. In our study, the mean oxygen saturation in the once-daily group decreased by 2% 1 day before exacerbation. In contrast, the oxygen saturation of the overnight group had decreased by 1.2% for several days before the onset of the exacerbation. The variation in the overnight group (±0.36%) was less than that in the once-daily group (±0.81%), which again would increase the potential to detect changes at exacerbation of COPD. However, it should be noted that these changes are small in magnitude and within the measurement error of the device.

Pulse Oximetry

We hypothesized that a continuous overnight monitoring approach might help in earlier detection of COPD exacerbations by eliminating non–exacerbation-related influences on the pulse oximetry signal, such as that from exercise, medication, and anxiety. We acknowledge that changes in heart rate and oxygen saturation with exercise may provide an alternative approach for the early detection of exacerbations, but our study was not designed to examine this.

Although our data suggested that monitoring pulse oximetry may help in early identification of exacerbations, there remains a need to prospectively test an algorithm that could be implemented in telehealth systems. We have shown that combining two variables in a combined oximetry score increases the potential to detect changes earlier, with sensitivity, specificity, and positive predictive value that would be clinically valuable. Our composite score crossed the significance threshold in both groups but was superior in the overnight group. The score in the overnight group crossed the threshold for 7 days before the treatment was initiated in the patient. The monitoring of heart rate and oxygen saturation in COPD telehealth services is widespread. In a cross-sectional survey on the use of telehealth in COPD, heart rate and oxygen saturation were the variables most commonly monitored, but there was no standardized method for detecting changes and thus defining alarm limits [26]. Consequently, tele-health users report a high frequency of false alarms. In a previous systematic review, we reported that the majority of telehealth studies were taking measurements intermittently [7]. This approach might not be optimal for using pulse oximetry to detect exacerbations as the signal might be affected by external factors such as exercise, medication, and anxiety. We have shown that overnight monitoring of heart rate and oxygen saturation permits the detection of changes earlier than once-daily monitoring.

Although symptoms changed as early as physiological variables, patients do not routinely quantify symptoms in relation to their own baseline, and therefore, objective monitoring has potential benefits in the context of both clinical practice and research (avoiding missing unreported exacerbations, increasing event rates, and thus, requiring fewer patients to meet an exacerbation end point).

We found that living alone and the presence of comorbidities may be associated with higher acceptance of overnight monitoring. Previous studies have shown patients' willingness to use home monitoring devices if they thought it would benefit their condition [27].

To our knowledge, this is the first study to monitor patients with COPD overnight to support earlier detection of exacerbations. We measured the pulse rate and oxygen saturation every 4 seconds, which reduced day-to-day variability in the measurement signal compared with once-daily monitoring. A significant limitation in our study was the higher-than-expected dropout rate. We believe this was because the equipment used in the overnight group was not designed for this purpose, patients lost interest in the study, and there was no real-time interaction with the health care provider with regard to their measurements. Although these factors are modifiable, we did not detect the number of exacerbations required to satisfy our power calculation, and therefore, the results are best considered as a pilot. Our findings have shown the potential for pulse oximetry to assist in the early detection of exacerbations, but we cannot draw a stronger conclusion about the efficacy because of the small number of exacerbations captured.

In conclusion, we have demonstrated that overnight pulse oximetry could facilitate earlier objective detection of COPD exacerbations. Overnight pulse oximetry allowed earlier detection compared with once-daily monitoring. Heart rate was superior to oxygen saturation and PEF and combining heart rate and oxygen saturation data provided the best performance.

Acknowledgments

The authors gratefully acknowledge the study participants and support they received from the pulmonary rehabilitation sites and staff.

Authors' Contributions

AR and JH conceived and designed the study, and the methodology was revised by SM, YA, AA, ML, JA, EP, and SQ. AR and JH performed the initial analysis, interpretation, and evaluation of the data. AR wrote the first manuscript draft, and all the authors revised it for important intellectual content. All the authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT checklist.
[PDF File (Adobe PDF File), 206 KB - mhealth_v8i11e17597_app1.pdf]

References

1. Burden of COPD. World Health Organization. URL: https://www.who.int/respiratory/copd/burden/en/ [accessed 2019-01-29] [WebCite Cache ID who.int/respiratory/copd/burden/en/]
2. World Health Statistics. World Health Organization. 2008. URL: https://www.who.int/whosis/whostat/EN_WHS08_Full.pdf?ua=1 [accessed 2019-01-18]
3. Lee S, Huang M, Kang J, Lin C, Park MJ, Oh Y, Investigators of the Predictive Ability of CAT in Acute Exacerbations of COPD (PACE) Study. The COPD assessment test (CAT) assists prediction of COPD exacerbations in high-risk patients. Respir Med 2014 Apr;108(4):600-608 [FREE Full text] [doi: 10.1016/j.rmed.2013.12.014] [Medline: 24456695]
4. Chronic Obstructive Pulmonary Disease (COPD) Facts Sheet. World Health Organization. 2017 Dec 01. URL: https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd) [accessed 2019-01-29] [WebCite Cache ID https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd)]
5. Wilkinson TM, Donaldson GC, Hurst JR, Seemungal TA, Wedzicha JA. Early therapy improves outcomes of exacerbations of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2004 Jun 15;169(12):1298-1303. [doi: 10.1164/rccm.200310-1443OC] [Medline: 14990395]
6. Seemungal TA, Donaldson GC, Bhowmik A, Jeffries DJ, Wedzicha JA. Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2000 May;161(5):1608-1613. [doi: 10.1164/ajrccm.161.5.9908022] [Medline: 10806163]
7. Al Rajeh A, Hurst J. Monitoring of physiological parameters to predict exacerbations of chronic obstructive pulmonary disease (COPD): a systematic review. J Clin Med 2016 Nov 25;5(12):- [FREE Full text] [doi: 10.3390/jcm5120108] [Medline: 27897995]
8. Struik FM, Sprooten RT, Kerstjens HA, Bladder G, Zijnen M, Asin J, et al. Nocturnal non-invasive ventilation in COPD patients with prolonged hypercapnia after ventilatory support for acute respiratory failure: a randomised, controlled, parallel-group study. Thorax 2014 Sep;69(9):826-834. [doi: 10.1136/thoraxjnl-2014-205126] [Medline: 24781217]
9. Hurst JR, Donaldson GC, Quint JK, Goldring JJ, Patel AR, Wedzicha JA. Domiciliary pulse-oximetry at exacerbation of chronic obstructive pulmonary disease: prospective pilot study. BMC Pulm Med 2010 Oct 20;10:52 [FREE Full text] [doi: 10.1186/1471-2466-10-52] [Medline: 20961450]

10. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991 Dec;14(6):540-545. [doi: 10.1093/sleep/14.6.540] [Medline: 1798888]
11. Nagappa M, Liao P, Wong J, Auckley D, Ramachandran SK, Memtsoudis S, et al. Validation of the STOP-bang questionnaire as a screening tool for obstructive sleep apnea among different populations: a systematic review and meta-analysis. PLoS One 2015;10(12):e0143697 [FREE Full text] [doi: 10.1371/journal.pone.0143697] [Medline: 26658438]
12. Operator's Manual. Nonin Medical: Pulse and Regional Oximeters, Capnographs. URL: https://www.nonin.com/wp-content/uploads/2018/10/Operators-Manual-3150-with-BLE.pdf [accessed 2020-04-01]
13. GO2 Model 9570 Fingertip Pulse Oximeter. Nonin Medical: Pulse and Regional Oximeters, Capnographs. 2019. URL: https://www.nonin.com/wp-content/uploads/Operators-Manual-Go2.pdf [accessed 2020-04-01]
14. Hurst JR, Vestbo J, Anzueto A, Locantore N, Müllerova H, Tal-Singer R, et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) Investigators: susceptibility to exacerbation in chronic obstructive pulmonary disease. N Engl J Med 2010 Sep 16;363(12):1128-1138. [doi: 10.1056/NEJMoa0909883] [Medline: 20843247]
15. Wardini R, Dajczman E, Yang N, Baltzan M, Préfontaine D, Stathatos M, et al. Using a virtual game system to innovate pulmonary rehabilitation: safety, adherence and enjoyment in severe chronic obstructive pulmonary disease. Can Respir J 2013;20(5):357-361 [FREE Full text] [doi: 10.1155/2013/563861] [Medline: 24093115]
16. The Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. Global Initiative for Chronic Obstructive Lung. 2020. URL: https://goldcopd.org/wp-content/uploads/2019/12/GOLD-2020-FINAL-ver1.2-03Dec19_WMV.pdf [accessed 2020-09-18] [WebCite Cache ID https://goldcopd.org/]
17. Aaron SD, Donaldson GC, Whitmore GA, Hurst JR, Ramsay T, Wedzicha JA. Time course and pattern of COPD exacerbation onset. Thorax 2012 Mar;67(3):238-243. [doi: 10.1136/thoraxjnl-2011-200768] [Medline: 22008189]
18. Kon SS, Canavan JL, Jones SE, Nolan CM, Clark AL, Dickson MJ, et al. Minimum clinically important difference for the COPD assessment test: a prospective analysis. Lancet Respir Med 2014 Mar;2(3):195-203 [FREE Full text] [doi: 10.1016/S2213-2600(14)70001-3] [Medline: 24621681]
19. Mackay AJ, Donaldson GC, Patel AR, Jones PW, Hurst JR, Wedzicha JA. Usefulness of the chronic obstructive pulmonary disease assessment test to evaluate severity of COPD exacerbations. Am J Respir Crit Care Med 2012 Jun 1;185(11):1218-1224. [doi: 10.1164/rccm.201110-1843OC] [Medline: 22281834]
20. Murata GH, Lium DJ, Busby HK, Kapsner CO. Precision and accuracy of self-measured peak expiratory flow rates in chronic obstructive pulmonary disease. South Med J 1998 Oct;91(10):919-924. [doi: 10.1097/00007611-199810000-00005] [Medline: 9786286]
21. Emerman CL, Cydulka RK. Use of peak expiratory flow rate in emergency department evaluation of acute exacerbation of chronic obstructive pulmonary disease. Ann Emerg Med 1996 Feb;27(2):159-163. [doi: 10.1016/s0196-0644(96)70317-7] [Medline: 8629745]
22. Seemungal TA, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 1998 May;157(5 Pt 1):1418-1422. [doi: 10.1164/ajrccm.157.5.9709032] [Medline: 9603117]
23. Burton C, Pinnock H, McKinstry B. Changes in telemonitored physiological variables and symptoms prior to exacerbations of chronic obstructive pulmonary disease. J Telemed Telecare 2015 Jan;21(1):29-36. [doi: 10.1177/1357633X14562733] [Medline: 25475218]
24. Martín-Lesende I, Orruño E, Bilbao A, Vergara I, Cairo MC, Bayón JC, et al. Impact of telemonitoring home care patients with heart failure or chronic lung disease from primary care on healthcare resource use (the TELBIL study randomised controlled trial). BMC Health Serv Res 2013 Mar 28;13:118 [FREE Full text] [doi: 10.1186/1472-6963-13-118] [Medline: 23537332]
25. Pedone C, Chiurco D, Scarlata S, Incalzi RA. Efficacy of multiparametric telemonitoring on respiratory outcomes in elderly people with COPD: a randomized controlled trial. BMC Health Serv Res 2013 Mar 6;13:82 [FREE Full text] [doi: 10.1186/1472-6963-13-82] [Medline: 23497109]
26. Al Rajeh A, Steiner MC, Aldabayan Y, Aldhahir A, Pickett E, Quaderi S, et al. Use, utility and methods of telehealth for patients with COPD in England and Wales: a healthcare provider survey. BMJ Open Respir Res 2019;6(1):e000345. [doi: 10.1136/bmjresp-2018-000345] [Medline: 30956795]
27. Frühauf J, Schwantzer G, Ambros-Rudolph CM, Weger W, Ahlgrimm-Siess V, Salmhofer W, et al. Pilot study on the acceptance of mobile teledermatology for the home monitoring of high-need patients with psoriasis. Australas J Dermatol 2012 Feb;53(1):41-46. [doi: 10.1111/j.1440-0960.2011.00852.x] [Medline: 22309330]

Abbreviations

ANOVA: analysis of variance
CAT: Chronic obstructive pulmonary disease assessment test
COPD: chronic obstructive pulmonary disease
FEV1: forced expiratory volume in one second
FVC: forced vital capacity

MCID: minimal clinically important difference
PEF: peak expiratory flow

Edited by G Eysenbach; submitted 23.12.19; peer-reviewed by J Parak, J Sedlák; comments to author 02.03.20; revised version received 17.04.20; accepted 15.05.20; published 13.11.20.

Please cite as:
Al Rajeh AM, Aldabayan YS, Aldhahir A, Pickett E, Quaderi S, Alqahtani JS, Mandal S, Lipman MCI, Hurst JR
Once Daily Versus Overnight and Symptom Versus Physiological Monitoring to Detect Exacerbations of Chronic Obstructive Pulmonary Disease: Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(11):e17597
URL: https://mhealth.jmir.org/2020/11/e17597
doi:10.2196/17597
PMID:33185560

Original Paper

Implementing Mobile Health–Enabled Integrated Care for Complex Chronic Patients: Patients and Professionals' Acceptability Study

Jordi de Batlle[1,2], PhD; Mireia Massip[1], BSc; Eloisa Vargiu[3], PhD; Nuria Nadal[4], MD; Araceli Fuentes[5], MD; Marta Ortega Bravo[6,7,8], MD; Jordi Colomina[9], MD; Reis Drudis[10], MD; Montserrat Torra[10], MD; Francesc Pallisó[9], PhD; Felip Miralles[3], PhD; Ferran Barbé[1,2], MD; Gerard Torres[1,2], MD; CONNECARE-Lleida Group[11]

[1]Group of Translational Research in Respiratory Medicine, Institut de Recerca Biomedica de Lleida, Lleida, Spain

[2]Center for Biomedical Network Research in Respiratory Diseases, Madrid, Spain

[3]eHealth Unit, Eurecat, Centre Tecnòlogic de Catalunya, Barcelona, Spain

[4]Gerència Territorial de Barcelona, Institut Català de la Salut, Barcelona, Spain

[5]Atenció Primària Àmbit Lleida, Lleida, Spain

[6]Research Support Unit Lleida, Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Lleida, Spain

[7]Centre d'Atenció Primària Cappont, Gerència Territorial de Lleida, Institut Català de la Salut, Lleida, Spain

[8]Universitat de Lleida, Lleida, Spain

[9]Servei de Cirurgia Ortopèdica i Traumatologia, Hospital de Santa Maria de Lleida, Lleida, Spain

[10]Unitat de Dolor Agut, Hospital de Santa Maria de Lleida, Lleida, Spain

[11]Institut de Recerca Biomedica de Lleida, Lleida, Spain

Corresponding Author:
Gerard Torres, MD
Group of Translational Research in Respiratory Medicine
Institut de Recerca Biomedica de Lleida
80 Rovira Roure Lleida
Lleida, 25198
Spain
Phone: 34 973705372
Email: gtorres@gss.cat

Abstract

Background: Integrated care (IC) can promote health and social care efficiency through prioritization of preventive patient-centered models and defragmentation of care and collaboration across health tiers, and mobile health (mHealth) can be the cornerstone allowing for the adoption of IC.

Objective: This study aims to assess the acceptability, usability, and satisfaction of an mHealth-enabled IC model for complex chronic patients in both patients and health professionals.

Methods: As part of the CONNECARE Horizon 2020 project, a prospective, pragmatic, 2-arm, parallel, hybrid effectiveness-implementation trial was conducted from July 2018 to August 2019 in a rural region of Catalonia, Spain. Home-dwelling patients 55 years and older with chronic conditions and a history of hospitalizations for chronic obstructive pulmonary disease or heart failure (use case [UC] 1), or a scheduled major elective hip or knee arthroplasty (UC2) were recruited. During the 3 months, patients experienced an mHealth-enabled IC model, including a self-management app for patients, a set of integrated sensors, and a web-based platform connecting professionals from different settings or usual care. The Person-Centered Coordinated Care Experience Questionnaire (P3CEQ) and the Nijmegen Continuity Questionnaire (NCQ) assessed person-centeredness and continuity of care. Acceptability was assessed for IC arm patients and staff with the Net Promoter Score (NPS) and the System Usability Scale (SUS).

Results: The analyses included 77 IC patients, 58 controls who completed the follow-up, and 30 health care professionals. The mean age was 78 (SD 9) years in both study arms. Perception of patient-centeredness was similarly high in both arms (usual care: mean P3CEQ score 16.1, SD 3.3; IC: mean P3CEQ score 16.3, SD 2.4). IC patients reported better continuity of care than controls (usual care: mean NCQ score 3.7, SD 0.9; IC: mean NCQ score 4.0, SD 1; *P*=.04). The scores for patient acceptability (UC1: NPS +67%; UC2: NPS +45%) and usability (UC1: mean SUS score 79, SD 14; UC2: mean SUS score 68, SD 24) were outstanding. Professionals' acceptability was low (UC1: NPS −25%; UC2: NPS −35%), whereas usability was average (UC1: mean SUS score

63, SD 20; UC2: mean SUS score 62, SD 19). The actual use of technology was high; 77% (58/75) of patients reported physical activity for at least 60 days, and the ratio of times reported over times prescribed for other sensors ranged from 37% for oxygen saturation to 67% for weight.

Conclusions: The mHealth-enabled IC model showed outstanding results from the patients' perspective in 2 different UCs but lacked maturity and integration with legacy systems to be fully accepted by professionals. This paper provides useful lessons learned through the development and assessment process and may be of use to organizations willing to develop or implement mHealth-enabled IC for older adults.

(JMIR Mhealth Uhealth 2020;8(11):e22136) doi:10.2196/22136

KEYWORDS

mHealth; eHealth; patient acceptance of health care; patient satisfaction; health plan implementation; chronic diseases

Introduction

In recent decades, socioeconomic development has increased life expectancy and led to progressively aging populations with an increased burden of chronic diseases [1]. This increased disease burden has heavily affected the already overburdened health and social care systems, which struggle to provide adequate services with limited resources. Traditional care models are falling short in adequately responding to the needs of chronic patients. The struggle of different care settings to communicate efficiently among themselves leads to care fragmentation and patients being left with the feeling of starting anew after every transition [2]. Moreover, patients are often passive actors without the necessary empowerment and no clear role in their own management [2]. Overall, there is a need for a profound redesign of how care is provided to older people with chronic health conditions, with a focus on patient participation, care quality, and system sustainability [3]. Integrated care (IC) models were created to address these challenges, with the aim of generating efficiencies through the adoption of patient-centered models, promotion of efficient continuity of care across settings, and prioritization of preventive strategies [4]. eHealth and mobile health (mHealth) can be the cornerstone allowing for the adoption of IC models [5].

In this scenario, the Horizon 2020 European Union Research and Innovation project CONNECARE—Personalized Connected Care for Complex Chronic Patients—attempted to co-design, develop, deploy, and evaluate a smart adaptive IC model for complex chronic patients (CCPs) [6]. From April 2016 to December 2019, participants in CONNECARE co-designed and experienced an organizational model supported by an eHealth platform that allows IC. The IC model allowed shifting from a conventional reactive care to a home-based preventive model built on cross-setting collaboration and interoperability, patient empowerment, and health risk prediction and management based on the analysis of patient-specific and population-based data. The model was supported by an advanced eHealth platform, based on information and communication technologies and Internet of Things, offering a cross-setting web-based Smart Adaptive Case Management (SACM) system for professionals and an mHealth self-management app with 3-level monitoring for patients. Finally, the CONNECARE IC model and supporting eHealth platform were tailored to different settings involving CCPs across Europe, including Lleida and Barcelona in Spain, Groningen in the Netherlands, and Ashdod in Israel.

Most novel IC interventions are assessed in light of the Triple Aim compass, which enhances patient experience, improves population health, and reduces overall costs [7]. According to the Triple Aim concept, patients' acceptability and satisfaction are key aspects for the large-scale adoption of novel management strategies. However, it is equally important that the involved health and social care professionals feel that any change in their routines will allow them to provide better care to their patients. This key concept has been described well in the Quadruple Aim, which expands Triple Aim to include the improvement of the work life of clinicians and staff [8]. Therefore, the assessment of acceptability and satisfaction in patients and professionals should be an unavoidable aspect in the evaluation of novel IC strategies, complementing effectiveness and cost-effectiveness assessments.

The CONNECARE mHealth-enabled IC model is being deployed and tested in different European settings. This paper focuses on the assessment of acceptability and satisfaction in patients and professionals in relation to the implementation of mHealth-enabled IC in the rural region of Lleida—Catalonia (Spain)—which had historically struggled with low levels of cross-setting interoperability, limiting the capacity of professionals from different settings to provide a coordinated response to patients' needs, and limited patient empowerment, with patients having mostly passive roles throughout their care paths.

Methods

Study Design

A prospective, pragmatic, 2-arm, parallel, type 1 hybrid effectiveness-implementation trial [9] that assesses patients and professionals' acceptability of a 3-month mHealth-enabled IC intervention as compared with that of usual care was conducted. The study was conducted from July 2018 to August 2019 in Lleida, which is a large rural area of more than 4300 km^2, including 2 tertiary hospitals, University Hospital Arnau de Vilanova and University Hospital Santa Maria, and a network of 23 primary care centers spread across the whole territory, providing services to 400,000 citizens.

Target Population

The intervention was deployed for 2 different use cases (UCs): (1) home-dwelling patients, with chronic conditions and a history of visits to the emergency room (ER) leading to hospitalizations (UC1); and (2) home-dwelling patients, with chronic conditions, undergoing a major elective hip or knee arthroplasty surgery (UC2). The specific eligibility criteria included the following: age more than 55 years, having a hospital admission because of a respiratory or cardiovascular event (UC1), having a programmed major elective hip or knee arthroplasty surgery (UC2), not having dementia or cognitive impairment (Global Deterioration Scale<5 [10]), LACE index for readmission score>7 (UC1) [11], American Society of Anesthesiologists Physical Status Classification System II or III (UC2) [12], being assigned to a primary care center in the region, living at home and being discharged back to the community, and passing a basic technological test assessing home connectivity and patients and/or care givers' competence with the use of technology (Multimedia Appendix 1 [13,14]).

Recruitment

Patients were recruited either during an unanticipated admission to the hospital through the ER (UC1) or at the time of surgery (UC2). All patients were identified based on data from electronic medical records (EMRs) and contacted by a case manager. Regardless of the UC, once a patient was recruited for the intervention arm, the search for a similar control began. All patients and their caregivers, regardless of the UC and study arm, received a face-to-face explanation about the study.

Intervention

Patients in the intervention arm experienced an mHealth-enabled IC model, including (1) a preliminary assessment of the patient's health status using several questionnaires, tests, and indices specific to their main chronic diseases and social needs done before hospital discharge for UC1 patients and at the time of scheduled surgery for UC2 patients; (2) access to a self-management app with status and performance reports, a virtual coach with customizable automated feedback, and full communication with the care team and guidance on its day-to-day use, taking into account that the app could be managed directly by the patients or indirectly by the informal caregivers or relatives; (3) a Fitbit Flex 2 (Fitbit) digital activity tracker [15] and additional sensors deemed necessary by the care team [16]: digital pulse-oximeter, digital scale, and digital blood pressure monitor, all of them fully integrated into the self-management app; (4) a patient's profile in the SACM web-based platform, which would be accessible by all the involved professionals (hospital, primary, and social care) and used to coordinate and communicate with professionals in the different settings, control the patient's evolution, and contact the patient if needed; and (5) assignment of a case manager in charge of supervising the whole process and being the main patient contact point. Additional details on the IC model and the required implementation efforts can be found in Multimedia Appendix 1. It must be noted that, as part of the CONNECARE project, the supporting technology used by patients and professionals was being developed and fine-tuned throughout the study. Usual care arm patients were managed from primary care.

Data Collection

Patients' characteristics were collected at recruitment, including age, sex, main chronic diseases, Charlson Comorbidity Index [17], Barthel Index for Activities of Daily Living [18], and the Pfeiffer Mental Status Questionnaire [19]. The main patient outcomes were collected after 3 months and included patient's perception of person-centeredness, assessed by the Person-Centered Coordinated Care Experience Questionnaire (P3CEQ) [20]; patient's perception of continuity of care, assessed by questions G1 to G5 of the Nijmegen Continuity Questionnaire (NCQ) [21]; satisfaction with the IC platform in IC arm patients and staff, assessed by the Net Promoter Score (NPS) [22] and the System Usability Scale (SUS) [23]; and the actual use of the different elements of the IC platform by patients. The NPS was based on the question "How likely is it that you would recommend our system CONNECARE to a family member or friend?" to be answered in a 0 (*not at all likely*) to 10 (*extremely likely*) scale. Individuals scoring 9 or 10 were considered as *promoters*, individuals scoring 7 or 8 as *passives*, and individuals scoring 0 to 6 as *detractors*. The final NPS score was obtained by subtracting the proportion of *detractors* from the proportion of *promoters*, and it could range from −100% to +100% (a positive score is considered good, +50% is considered excellent, and anything more than +70% is exceptional) [24].

Statistical Analyses

Participants' baseline characteristics were described by the number (percentage), mean (SD), or median (P25-P75), as appropriate. Comparisons between IC and usual care patients' baseline characteristics were performed using chi-square test, *t* test, or Kruskal-Wallis test, as appropriate. Comparisons between IC and usual care patients' person-centeredness and continuity of care were performed using chi-square test or *t* test, as appropriate, excluding patients answering "Don't know" or "No answer." Satisfaction with the IC platform in IC patients and staff were described using mean (SD) or median (P25-P75), as appropriate. Finally, the actual use by patients of the different elements of the IC platform was described by reporting the proportion of times reported over times prescribed, using median (P25-P75). Data analyses were conducted using Stata, version 12.1 (StataCorp). The threshold for significance was set at .05.

Implementation Framework

The Consolidated Framework for Implementation Research (CFIR) [13] was used to assess implementation aspects. A detailed description of the implementation strategies and framework can be found in Multimedia Appendix 1.

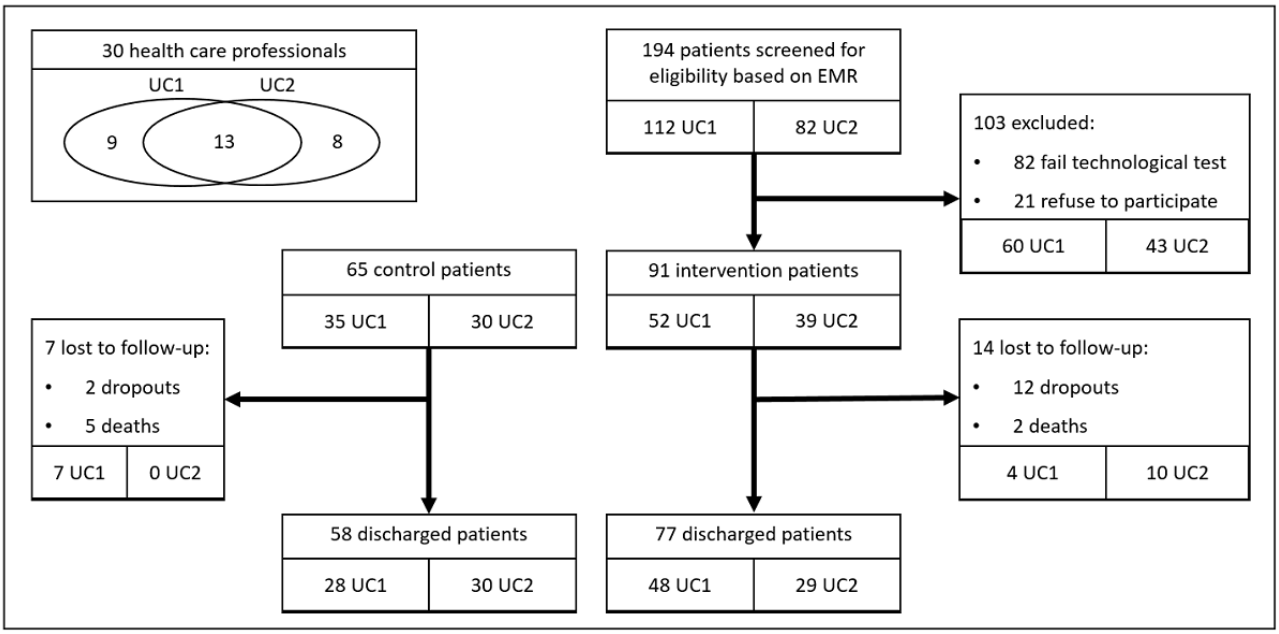
Ethical Considerations

This study was approved by the Ethics Committee of Hospital Arnau de Vilanova (CEIC-1685), and all participants provided written informed consent. All collected data were handled and stored in accordance with current National and International legislation.

Results

In this study, up to 194 patients were screened for eligibility (112 for UC1 and 82 for UC2). After excluding patients who did not meet the inclusion criteria, 91 patients were recruited for the intervention arm and 65 for the usual care arm. Final analyses were based on 77 IC patients and 58 usual care control patients who completed the follow-up and 30 health care professionals (Figure 1).

Figure 1. Study flowchart. EMR: electronic medical record; UC: use case.



The main characteristics of the patients included in the study are shown in Table 1. The mean age was 78 (SD 9) years in both study arms, with patients in UC1 being 10 years older than patients in UC2 on average. Patients in UC1 had a higher burden of comorbidities than patients in UC2 (Charlson Comorbidity Index in UC1: mean 6.9, SD 2.1; UC2: mean 4.2, SD 1.6; P<.001). No statistically significant differences were found between patients in the usual care arm and patients in the IC arm, regardless of UC. Regarding the analyses of health care professionals, UC1 included 1 hospital case manager, 3 hospital physicians, 3 primary care case managers, 9 primary care physicians, and 6 primary care nurses, and UC2 included 1 hospital case manager, 1 hospital physician, 1 hospital surgeon, 1 hospital anesthesiologist, 2 primary care case managers, 9 primary care physicians, and 6 primary care nurses.

Table 1. Baseline characteristics of patients in usual care and integrated care.

Characteristics	Use case 1			Use case 2		
	Usual care (n=28)	Integrated care (n=48)	P[a] value	Usual care (n=30)	Integrated care (n=29)	P[a] value
Sex (male), n (%)	17 (61)	24 (50)	.37	8 (27)	12 (41)	.23
Age (years), mean (SD)	82 (8)	82 (7)	.88	73 (8)	72 (9)	.50
Charlson[b] (score), mean (SD)	7.4 (2.1)	6.7 (2.0)	.15	4.2 (1.7)	4.2 (1.5)	.88
Barthel[c] (score), median (P25-P75)	90 (73-95)	90 (68-100)	.40	95 (90-100)	100 (95-100)	.16
Pfeiffer[d], NI[e], n (%)	21 (75)	37 (77)	.67	30 (100)	27 (93)	.14

[a]Chi-square test, *t* test, or Kruskal-Wallis equality-of-populations rank test, as appropriate.

[b]Charlson Comorbidity Index.

[c]Barthel Index for Activities of Daily Living.

[d]Pfeiffer Mental Status Questionnaire.

[e]NI: no impairment.

Regardless of UC and intervention arm, the perception of patient-centeredness was very high, with a mean P3CEQ score of 16.1 (SD 3.3) in the usual care arm and 16.3 (SD 2.4) in the IC arm. Regarding the continuity of care, patients in IC scored better than patients in usual care (usual care: mean NCQ G1-G5 score 3.7, SD 0.9; integrated care: mean NCQ G1-G5 score 4.0, SD 1.0; P=.04). Further information on patient-perceived person-centeredness and continuity of care can be found in Multimedia Appendix 1.

Patients and professionals' satisfaction with the deployed technology for IC is presented in Tables 2 and 3. Patients in UC1 reported an overall NPS score of +67% and patients in UC2 of +45%, whereas professionals involved in UC1 scored −25% and in UC2 −35%. Similarly, although the mean score for SUS was 75 (SD 19), the mean score for professionals was 61 (SD 20).

Table 2. Integrated care patients' satisfaction with the technology.

Measures	All (n=77)	UC[a]1 (n=48)	UC2 (n=29)
NPS[b] questions (0 [poor] to 10 [good]), median (P25-P75)			
Overall satisfaction	10 (8-10)	10 (8-10)	10 (8-10)
Easiness of use	9 (8-10)	9 (8-10)	8 (5-10)
Ability to be used without help	9 (7-10)	9 (8-10)	8 (5-10)
Would you recommend it	10 (8-10)	10 (8-10)	9 (8-10)
NPS score (−100% to +100%)	+58%	+67%	+45%
SUS[c] score (0 [awful] to 100 [excellent])			
Mean (SD)	75 (19)	79 (14)	68 (24)
Users scoring over 68 points, n (%)	54 (70)	38 (79)	16 (55)

[a]UC: use case.

[b]NPS: Net Promoter Score.

[c]SUS: System Usability Scale.

Table 3. Staff's satisfaction with the technology.

Measures	All (n=30)	UC[a]1 (n=22)	UC2 (n=21)
NPS[b] questions (0 [poor] to 10 [good]), median (P25-P75)			
Overall satisfaction	6 (5-8)	6 (5-8.5)	6.5 (5-8)
Easiness of use	6 (6-7)	6.5 (5-8)	6 (4.5-7.5)
Ability to be used without help	6 (5-9)	6.5 (5-9)	6 (5.5-9)
Would you recommend it	6.5 (5-8)	6.5 (5.5-8.5)	6.5 (5-7.5)
NPS score (−100% to +100%):	−29%	−25%	−35%
SUS[c] score (0 [awful] to 100 [excellent])			
Mean (SD)	61 (20)	63 (20)	62 (19)
Users scoring over 68 points, n (%)	14 (47)	10 (46)	9 (43)

[a]UC: use case.

[b]NPS: Net Promoter Score.

[c]SUS: System Usability Scale.

Table 4 reports on the actual use of technology by the patients experiencing IC. The actual use of the physical activity tracker was outstanding, with up to 77% (58/75) of patients having reported measures over 60 days out of 90. The ratio of times reported over times prescribed for the rest of sensors that could be proposed to patients ranged from 37% for oxygen saturation to 67% for weight. Finally, the use of the messaging function allowing patients to ask or answer requirements for or from the professional care team was high, with a median (P25-P75) of 19 (10.5-41) in UC1 and 10 (5-22) in UC2 over the 90-day intervention.

Finally, the results of the evaluation of the implementation according to CFIR can be found in Multimedia Appendix 1.

Table 4. Integrated care patients' use of technology.

Technologies	All (n=77)	UC[a]1 (n=48)	UC2 (n=29)
Daily steps (Fitbit)			
Days reported, median (P25-P75)	86 (63-92)	86 (63-92)	87 (68.5-92)
Users reporting ≥60 days, n (%)	58 (77)	36 (77)	22 (79)
Weight (Withings), median (P25-P75)			
Times R/P[b]	67 (42-91)	67 (42-91)	NU[c]
Blood pressure (Withings), median (P25-P75)			
Times R/P	41 (32-50)	43 (37-50)	38 (21-58)
Heart rate (Withings), median (P25-P75)			
Times R/P	42 (36-51)	42 (36-51)	NU
Oxygen saturation (SpO_2; Withings), median (P25-P75)			
Times R/P	37 (22-42)	37 (22-42)	NU
Body temperature (Withings), median (P25-P75)			
Times R/P	41 (36-48)	NU	41 (36-48)
Messages to the care team, median (P25-P75)			
Total number	18 (7-37)	19 (10.5-41)	10 (5-22)

[a]UC: use case.

[b]R/P: reported or prescribed.

[c]NU: not used in the UC.

Discussion

Principal Findings

The assessment of the patients and professionals' acceptability of an mHealth-enabled IC program targeting home-dwelling CCP patients with a history of hospitalizations (UC1) or undergoing a major elective hip or knee arthroplasty surgery (UC2) showed 2 different perceptions. Although patients and/or informal caregivers or relatives reported a very high acceptability of the IC program and its supporting technology, professionals rated it as moderately poor. Patients reported very high perceptions of patient-centeredness, continuity of care, satisfaction, and usability of the IC platform, and matching with these positive perceptions, actual use of the different features of the IC platform was high. Although patients positively qualified their user experience [25], professionals felt the burden of a system under constant development, which in turn limited their experience and translated to moderately poor satisfaction with the IC platform.

Strengths and Limitations

This study has several strengths: (1) from day 1, an effort was made to involve all actors from different organizations who would participate in a large-scale deployment of the mHealth-supported IC program, which is key as the lack of cooperation between organizations and professionals is a well-known barrier for the implementation of IC [4]; (2) the involvement of informal carers as actors being important in facilitating the use of the patient's app in older patients; (3) the prescription and monitoring of patients' physical activity, as mobility impairment is frequent in people older than 65 years [26]; and (4) the geography of the implementation area, a large rural region of more than 4300 km^2, which could benefit the most from community-based integrated care initiatives that precluded unnecessary visits to primary care centers or hospitals. Similarly, there were several limitations: (1) the IC platform was in a constant process of refinement and addition of new functionalities; thus, the user experience was richer by the end of the implementation study compared with the very beginning; and (2) having a single entry point to the IC program, which was the hospital either after an ER admission (UC1) or at the time of surgery (UC2), as it is important that system-wide cross-organizational care pathways consider multiple entry points [27]. However, primary care centers are currently being considered as a potential additional entry point if the IC system is further implemented in the region.

Patients' Perspective: Comparison With Previous Work

The use of mHealth apps for patients with chronic conditions has been explored in the last decade, showing the potential for appropriate security level, effective monitoring, self-management, and communication [28]. However, a 2016 review of Apps for Heart Failure symptom monitoring and self-care reported that a minority of the available apps had the required quality, content, or functionality [29]. Our IC platform aimed to go beyond a patient's app and established a comprehensive IC model, including a patient's app, a portfolio of different sensors linked with the app, and a professional's web-based platform for monitoring, communication with the patient, and collaboration among professionals in different health settings. The first indicator of the success of an mHealth

intervention is the rate of dropouts. In our study, only 13% of patients in the IC arm decided to abandon the program. For instance, Bentley et al [30] reported that half of the participants in an mHealth-based self-management intervention for chronic obstructive pulmonary disease (COPD) withdrew from the study. The scores for acceptability (UC1: NPS +67%; UC2: NPS +45%) and usability (UC1: mean SUS score 79, SD 14; UC2: mean SUS score 68, SD 24) were outstanding. These results are better than most results obtained in existing mHealth tools targeting chronic patients, for example, mHealth tools aiming to improve quality of life in breast cancer survivors such as BENECA (NPS +7%) [31] or Oncokompas (NPS −36%) [32], an Intelligent Virtual Assistant for promoting behavior change and self-care in older people with type 2 diabetes (mean SUS 74, SD 13) [33], or a gait-monitoring mobile phone app for older users (mean SUS 60, SD 11) [34]. Similarly, interventions targeting surgical patients by means of education and communication apps have obtained positive results in terms of acceptability and usability [35,36], comparable with results obtained with automated phone messaging platforms [37,38]. Finally, preliminary results of variants of the CONNECARE IC model implemented in other European settings have been positive. For instance, the CONNECARE self-management app was rated positively by older patients with cancer who were offered remote home monitoring after surgery (mean SUS 74, SD 19; NPS +29%) [39].

Patients' Perspective: Lessons Learned

The key factors for the success of our comprehensive IC model from the patients' perspective have been (1) including a comprehensive set of features with the patient's app acting as a hub of services including the integration of monitoring devices, in line with a 2016 review of Apps for Heart Failure symptom monitoring and self-care reporting that a minority of the available apps had the required content or functionality [29]; (2) the involvement of patients since early phases of development, as proposed by Lundell et al [40] in a recent qualitative analysis of the use of home telemonitoring in patients with COPD; (3) the flexibility of potential end users, as the app could be managed directly by the patients (most UC2 patients) or by a relative or informal carer (most UC1 patients); (4) enabled bidirectional communication with the care team, potentially avoiding unnecessary visits to primary care centers or hospital; (5) appropriate feedback on the daily monitoring and patients' achieved goals, including personalized motivational advice; (6) push-up notifications to remind key events, tasks, or goals; and (7) ease of use and quality-of-life features, such as being translated to the different official languages in the region (Catalan and Spanish) or having several display settings including font size, as difficulty in using the technology is a common reason for withdrawal [30].

Professionals' Perspective

According to the Quadruple Aim, the improvement of the work life of clinicians and staff is a key factor for the adoption of new health programs [8]. In this study, the scores for professionals' acceptability were low (UC1: NPS −25%; UC2: NPS −35%), whereas scores for usability were moderately high (UC1: mean SUS score 63, SD 20; UC2: mean SUS score 62, SD 19). However, these ratings were directly related to the temporal constraints of the study setting, as professionals were required to use a system in a dynamic development and implementation process rather than a fully developed one. However, having the opportunity to directly participate in the development of the IC model and platform allowed professionals to feel engaged and propose changes and new features to be developed, which ultimately resulted in great engagement (no professionals dropped out of the implementation study). Understanding the factors influencing professionals' adoption of eHealth is complex [41], but the ability to provide quality care is key [42]. On the one hand, the potential of the IC platform to provide quality care was the key to professionals' engagement; being able to monitor key aspects of chronic diseases or monitoring pain after surgery, enabling communication between carers in different care settings; having the option to prescribe and monitor physical activity; or case managers having access to a geographical representation of patients in a map, with the possibility of selecting patients based on predefined characteristics or generating optimized routes for home visits, showed professionals the game-changing features of the platform. On the other hand, a system under constant development, not achieving a full integration with legacy EMR systems and the coexistence of 2 management systems (usual care and IC) at the same time (which implied some duplicity of tasks) were the main barriers to adoption.

Challenges for Large-Scale Deployment

Although usability and acceptability are key for the adoption of mHealth-enabled IC, large-scale adoption requires cost-effectiveness and an adequate reimbursement and payment model. On the one hand, regarding cost-effectiveness, the implementation of our IC model reduced unplanned contacts with the health system, reduced health costs, and was cost-effective, as reported elsewhere. On the other hand, designing of an imbursement and payment model capable of accommodating the costs of new roles and required technologies, while fully benefiting from the savings in terms of reductions in the use of health and social care resources, can be challenging, especially when the model involves different organizations and providers. Therefore, a firm positioning of the involved health authorities and governing bodies is required to fully fulfill the ambition of our mHealth-enabled IC model. The use of the CFIR framework highlighted the key barriers and facilitators for large-scale adoption (Multimedia Appendix 1).

Conclusions

The assessment of the patients and professionals' acceptability of an mHealth-enabled IC program showed outstanding results from the patients' perspective. However, the web-based professionals' platform needs to be fully matured and fully integrated into legacy systems before moving forward toward large-scale deployment. This paper, thus, provides useful lessons learned through the development and assessment process and may be of use for organizations willing to develop or implement mHealth-enabled IC for older adults.

Acknowledgments

The authors would like to express their sincerest gratitude to all the participating patients and their caregivers as well as all the involved health and social care professionals. All of them were key in the co-design of the CONNECARE integrated care model and its initial implementation in the region of Lleida.

The CONNECARE-Lleida group: Maria Aguilà Balastegui, Sandra Alexandre Loxano, Laila Al-Jouja Llorente, Tomás Alonso Sancho, Enrique Aparicio Bañeres, Ana Arce Vila, Jose Maria Baron Burriel, Ramon Bascompte Claret, Albert Bigorda Sague, Emilia Blanco Ponce, Maria Boldú Franque, Àngels Bosch Roig, Carmen Bravo Santiago, Alba Capdevila Sarramona, Aida Castelló Corretge, Jordi Colomina Morales, Montse Coma Gassó, Fina Cregenzan Ortiz, Dolors Del Pozo Garcia, Reis Drudis Morell, Mireia Falguera Vilamajó, Pere Farre Pagés, Yolanda Fauria Garcia, Anabel Fusalba Canales, Jara Gayan Ordas, Sergi Godia Lopez, Irene Gomez Companys, Jessica Gonzàlez Gutierrez, Anna Gort Oromí, Carme Jorge Tufet, Mercé Lavega Llorens, Laia Llort Samsó, Maria Rosa Lopez Cervelló, Belen Malla Clua, Josep Maria Marsol Mas, Teresita Martí Ribes, Diana Martin Capella, José Maria Martínez Barriuso, Esther Mateus Solé, Ramon Mazana Novellon, Petra Merino De los Santos, Miquel Mesas Julio, Sonia Minguet Vidal, Nuria Moles Porta, Luis Miguel Montaña Esteban, Dolors Morera Roset, Meritxell Moyà Oro, Irene Muñoz Del Campo, Francisco Nicolás Sánchez, Inés Ortiz Catalán, Mireia Ortiz Valls, Sonia Ortiz Congost, Jose Maria Palacin Peruga, Francesc Pallisó Folch, Eugeni Paredes Costa, Pablo Pastor Pueyo, Ana Pérez Sainz, Antonio Plana Blanco, Anna Planas Hiraldo, Pepita Pont Aldoma, Marife Quelle Alonso, Rebeca Ramirez Molinero, Maria Àngels Revés Juanbaro, Anna Ribé Miró, Eva Ribó Caubet, Rebeca Rodriguez Corbaton, Marina Rué Florensa, Oscar Sacristán García, Irene Sanmartí Forns, Maria Cruz Sanz Martinez, Neus Sendra Bordes, Montse Torra Riera, Maria Cruz Urgelés Castillón, Laia Utrillo Montagut, and Montse Vidal Ballesté.

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme (under grant agreement no. GA-689802). JB acknowledges receiving financial support from the Catalan Health Departament (Pla Estratègic de Recerca i Innovació en Salut 2016: SLT002/16/00364) and Instituto de Salud Carlos III (ISCIII; Miguel Servet 2019: CP19/00108), cofunded by the European Social Fund (ESF), *Investing in your future*, and the European Regional Development Fund, "A way to make Europe." The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Authors' Contributions

JB, EV, NN, FP, FM, FB, and GT participated in the conceptualization of the project. MM, EV, AF, MOB, JC, RD, MTR, and GT conducted data collection. MM, EV, and JB participated in data curation. JB conducted all statistical analyses. JB wrote the original draft of the manuscript. All authors reviewed the final manuscript. JB, EV, FM, FB, and GT secured funding for the project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Implementing mobile health–enabled integrated care for complex chronic patients: a patients and professionals' acceptability study.
[DOCX File , 1576 KB - mhealth_v8i11e22136_app1.docx]

References

1. World Report on Ageing and Health. Geneva, Switzerland: World Health Organization; 2015.
2. Harvey G, Dollard J, Marshall A, Mittinty MM. Achieving integrated care for older people: shuffling the deckchairs or making the system watertight for the future? Int J Health Policy Manag 2018 Apr 1;7(4):290-293 [FREE Full text] [doi: 10.15171/ijhpm.2017.144] [Medline: 29626395]
3. Integrated Care for Older People: Guidelines on Community-level Interventions to Manage Declines in Intrinsic Capacity. Geneva, Switzerland: World Health Organization; 2017.
4. Dates M, Lennox-chhugani N, Sant H, Pereira A, Tedeschi M. Health system performance assessment – integrated care assessment (20157303 HSPA). Luxembourg 2018:- [FREE Full text] [doi: 10.2875/627462]
5. Chouvarda IG, Goulis DG, Lambrinoudaki I, Maglaveras N. Connected health and integrated care: toward new models for chronic disease management. Maturitas 2015 Sep;82(1):22-27. [doi: 10.1016/j.maturitas.2015.03.015] [Medline: 25891502]
6. The CONNECARE Project. URL: http://www.connecare.eu/ [accessed 2020-03-19]
7. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. Health Aff (Millwood) 2008;27(3):759-769. [doi: 10.1377/hlthaff.27.3.759] [Medline: 18474969]

8. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12(6):573-576. [doi: 10.1370/afm.1713] [Medline: 25384822]
9. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
10. Reisberg B, Ferris S, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. Am J Psychiatry 1982 Sep;139(9):1136-1139. [doi: 10.1176/ajp.139.9.1136] [Medline: 7114305]
11. van Walraven C, Dhalla IA, Bell C, Etchells E, Stiell IG, Zarnke K, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. Can Med Assoc J 2010 Apr 6;182(6):551-557. [doi: 10.1503/cmaj.091117] [Medline: 20194559]
12. Sankar A, Johnson S, Beattie W, Tait G, Wijeysundera D. Reliability of the American Society of Anesthesiologists physical status scale in clinical practice. Br J Anaesth 2014 Sep;113(3):424-432 [FREE Full text] [doi: 10.1093/bja/aeu100] [Medline: 24727705]
13. Consolidated Framework for Implementation Research (CFIR). URL: https://cfirguide.org/constructs/ [accessed 2020-08-25]
14. ACT academy. Plan, Do, Study, Act (PDSA) cycles and the model for improvement. NHS improvement. 2018. URL: https://improvement.nhs.uk/documents/2142/plan-do-study-act.pdf [accessed 2020-08-25]
15. Fitbit Flex 2 Digital Activity Tracker. URL: http://www.fitbit.com/in/flex2 [accessed 2020-08-25]
16. Withings. URL: http://www.withings.com/ [accessed 2020-08-25]
17. Charlson ME, Pompei P, Ales KL, MacKenzie C. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
18. Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud 1988;10(2):61-63. [doi: 10.3109/09638288809164103] [Medline: 3403500]
19. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. J Am Geriatr Soc 1975 Oct;23(10):433-441. [doi: 10.1111/j.1532-5415.1975.tb00927.x] [Medline: 1159263]
20. Lloyd H, Fosh B, Whalley B, Byng R, Close J. Validation of the person-centred coordinated care experience questionnaire (P3CEQ). Int J Qual Health Care 2019 Aug 1;31(7):506-512 [FREE Full text] [doi: 10.1093/intqhc/mzy212] [Medline: 30508089]
21. Uijen AA, Schers HJ, Schellevis FG, Mokkink HG, van Weel C, van den Bosch WJ. Measuring continuity of care: psychometric properties of the Nijmegen Continuity Questionnaire. Br J Gen Pract 2012 Jul;62(600):e949-e957. [doi: 10.3399/bjgp12X652364] [Medline: 22782001]
22. Reichheld F. The one number you need to grow. Harv Bus Rev 2003 Dec;81(12):46-54, 124. [Medline: 14712543]
23. Hell B, Meiert J. SUS - A Quick and Dirty Usability Scale. URL: https://hell.meiert.org/core/pdf/sus.pdf [accessed 2020-06-12]
24. Hamilton DF, Lane JV, Gaston P, Patton JT, Macdonald DJ, Simpson AH, et al. Assessing treatment outcomes using a single question: the net promoter score. Bone Joint J 2014 May;96-B(5):622-628. [doi: 10.1302/0301-620X.96B5.32434] [Medline: 24788496]
25. Vargiu E, Dorado A, Massip M, Gómez-Martínez M, Torres G, de Batlle J, et al. User Experience in CONNECARE: Results and Lessons Learnt in Lleida. XPA Barcelona. 2019. URL: http://www.xpabcn.com/revista/index.php/XPAHC/article/view/22 [accessed 2020-10-10]
26. Mitra S, Sambamoorthi U. Disability prevalence among adults: estimates for 54 countries and progress toward a global estimate. Disabil Rehabil 2014;36(11):940-947. [doi: 10.3109/09638288.2013.825333] [Medline: 23962193]
27. Borgermans L, Devroey D. A policy guide on integrated care (PGIC): lessons learned from EU project integrate and beyond. Int J Integr Care 2017 Sep 25;17(4):8 [FREE Full text] [doi: 10.5334/ijic.3295] [Medline: 29588631]
28. Wang J, Wang Y, Wei C, Yao N, Yuan A, Shan Y, et al. Smartphone interventions for long-term health management of chronic diseases: an integrative review. Telemed J E Health 2014 Jun;20(6):570-583. [doi: 10.1089/tmj.2013.0243] [Medline: 24787747]
29. 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]
30. Bentley CL, Powell L, Potter S, Parker J, Mountain GA, Bartlett YK, et al. The use of a smartphone app and an activity tracker to promote physical activity in the management of chronic obstructive pulmonary disease: randomized controlled feasibility study. JMIR Mhealth Uhealth 2020 Jun 3;8(6):e16203 [FREE Full text] [doi: 10.2196/16203] [Medline: 32490838]
31. Lozano-Lozano M, Cantarero-Villanueva I, Martin-Martin L, Galiano-Castillo N, Sanchez M, Fernández-Lao C, et al. A mobile system to improve quality of life via energy balance in breast cancer survivors (Beneca mhealth): prospective test-retest quasiexperimental feasibility study. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14136 [FREE Full text] [doi: 10.2196/14136] [Medline: 31237570]

32. Melissant HC, Verdonck-de Leeuw IM, Lissenberg-Witte BI, Konings IR, Cuijpers P, Van Uden-Kraan CF. 'Oncokompas', a web-based self-management application to support patient activation and optimal supportive care: a feasibility study among breast cancer survivors. Acta Oncol 2018 Jul;57(7):924-934. [doi: 10.1080/0284186X.2018.1438654] [Medline: 29451059]
33. Balsa J, Félix I, Cláudio AP, Carmo MB, Silva IC, Guerreiro A, et al. Usability of an intelligent virtual assistant for promoting behavior change and self-care in older people with type 2 diabetes. J Med Syst 2020 Jun 13;44(7):130. [doi: 10.1007/s10916-020-01583-w] [Medline: 32533367]
34. Zhong R, Rau PP. A mobile phone-based gait assessment app for the elderly: development and evaluation. JMIR Mhealth Uhealth 2020 Feb 29;8(2):e14453 [FREE Full text] [doi: 10.2196/14453] [Medline: 32452821]
35. Morte K, Marenco C, Lammers D, Bingham J, Sohn V, Eckert M. Utilization of mobile application improves perioperative education and patient satisfaction in general surgery patients. Am J Surg 2020 Apr 16 epub ahead of print. [doi: 10.1016/j.amjsurg.2020.03.034] [Medline: 32381263]
36. Gordon CR, Rezzadeh KS, Li A, Vardanian A, Zelken J, Shores JT, et al. Digital mobile technology facilitates HIPAA-sensitive perioperative messaging, improves physician-patient communication, and streamlines patient care. Patient Saf Surg 2015;9:21 [FREE Full text] [doi: 10.1186/s13037-015-0070-9] [Medline: 26136830]
37. 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]
38. 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]
39. Jonker LT, Plas M, de Bock GH, Buskens E, van Leeuwen BL, Lahr MM. Remote home monitoring of older surgical cancer patients: perspective on study implementation and feasibility. Ann Surg Oncol 2020 Jun 29 epub ahead of print. [doi: 10.1245/s10434-020-08705-1] [Medline: 32602060]
40. Lundell S, Modig M, Holmner A, Wadell K. Perceptions of home telemonitoring use among patients with chronic obstructive pulmonary disease: qualitative study. JMIR Mhealth Uhealth 2020 Jun 3;8(6):e16343 [FREE Full text] [doi: 10.2196/16343] [Medline: 32490844]
41. Gaveikaite V, Grundstrom C, Lourida K, Winter S, Priori R, Chouvarda I, et al. Developing a strategic understanding of telehealth service adoption for COPD care management: A causal loop analysis of healthcare professionals. PLoS One 2020;15(3):e0229619 [FREE Full text] [doi: 10.1371/journal.pone.0229619] [Medline: 32134958]
42. Friedberg M, Chen P, Van BK. Factors Affecting Physician Professional Satisfaction and Their Implications for Patient Care, Health Systems and Health Policy. Rand Corporation. 2013. URL: http://www.rand.org/content/dam/rand/pubs/research_reports/RR400/RR439/RAND_RR439.pdf [accessed 2020-06-18]

Abbreviations

CCP: complex chronic patient
CFIR: Consolidated Framework for Implementation Research
COPD: chronic obstructive pulmonary disease
EMR: electronic medical record
ER: emergency room
IC: integrated care
mHealth: mobile health
NCQ: Nijmegen Continuity Questionnaire
NPS: Net Promoter Score
P3CEQ: Person-Centered Coordinated Care Experience Questionnaire
SACM: Smart Adaptive Case Management
SUS: System Usability Scale
UC: use case

Edited by G Eysenbach; submitted 04.07.20; peer-reviewed by I Cano, G Harvey; comments to author 04.08.20; revised version received 31.08.20; accepted 12.09.20; published 20.11.20.

Please cite as:
de Batlle J, Massip M, Vargiu E, Nadal N, Fuentes A, Ortega Bravo M, Colomina J, Drudis R, Torra M, Pallisó F, Miralles F, Barbé F, Torres G, CONNECARE-Lleida Group
Implementing Mobile Health–Enabled Integrated Care for Complex Chronic Patients: Patients and Professionals' Acceptability Study
JMIR Mhealth Uhealth 2020;8(11):e22136
URL: https://mhealth.jmir.org/2020/11/e22136
doi:10.2196/22136
PMID:33216004

Original Paper

Internet and Telerehabilitation-Delivered Management of Rotator Cuff–Related Shoulder Pain (INTEL Trial): Randomized Controlled Pilot and Feasibility Trial

Peter Malliaras[1], BPhysio, PhD; Kate Cridland[1], BPhysio; Ruben Hopmans[1], BMM, MIT; Simon Ashton[1], BPhysio; Chris Littlewood[2], BPhysio, Phd; Richard Page[3], BMedSci, MBBS; Ian Harris[4,5,6], MBBS, PhD; Helen Skouteris[7], BSc, PhD; Terry Haines[8], BPhysio, PhD

[1]Physiotherapy Department, School of Primary and Allied Health Care, Faculty of Medicine Nursing and Health Science, Monash University, Melbourne, Australia

[2]Faculty of Health, Psychology and Social Care, Manchester Metropolitan University, Manchester, United Kingdom

[3]Barwon Orthopaedic Research and Education, Barwon Health and School of Medicine, Deakin University, Geelong, Australia

[4]Whitlam Orthopaedic Research Centre, Ingham Institute for Applied Medical Research, Sydney, Australia

[5]South Western Sydney Clinical School, University of New South Wales, Sydney, Australia

[6]Liverpool Hospital, Sydney, Australia

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

[8]School of Primary and Allied Health Care, Faculty of Medicine Nursing and Health Science, Monash University, Melbourne, Australia

Corresponding Author:
Peter Malliaras, BPhysio, PhD
Physiotherapy Department, School of Primary and Allied Health Care, Faculty of Medicine Nursing and Health Science
Monash University
Moorooduc Hwy, Frankston
Melbourne, 3199
Australia
Phone: 61 0400206480
Email: peter.malliaras@monash.edu

Abstract

Background: Rotator cuff–related shoulder pain (RCRSP) is a common and disabling musculoskeletal condition. Internet-based and telerehabilitation delivery of recommended care may improve access to care and improve adherence and outcomes.

Objective: The primary aim of this pilot randomized controlled trial was to assess the feasibility of a 12-week internet-delivered intervention for RCRSP comparing advice only, recommended care, and recommended care with group-based telerehabilitation.

Methods: Reporting was in accordance with the Consolidated Standards of Reporting Trials (CONSORT) checklist for pilot and feasibility trials. People with a primary complaint of RCRSP for 3 months or longer were identified via a paid Facebook strategy. Screening involved an online questionnaire followed by a 20-minute telehealth assessment. Participants were randomly allocated (via a Zelen design) to receive (1) advice only, (2) recommended care (internet-delivered evidence-based exercise and education), or (3) recommended care and telerehabilitation (including a weekly group teleconference session). Progression criteria for a full-scale trial included (1) recruitment of 20% or greater of eligible participants, (2) acceptable adherence (two or more of the three prescribed weekly sessions) among 70% or greater of participants, (3) 80% or greater retention of participants, (4) absence of intervention-related serious adverse events, and (5) 80% or greater response rates to questionnaires. Secondary clinical and patient knowledge outcomes were collected (via email or text) at baseline, six weeks, and 12 weeks (for clinical and patient knowledge), and within-group change was reported descriptively.

Results: We enrolled 36 of 38 (95%) eligible participants and all participants were recruited within a 3-week period. Of the 36 participants, 12 participants were allocated to each of the three trial arms. The mean age of participants was between 51 and 56 years, and 83% (10/12) to 92% (11/12) were female. Retention at the 12-week endpoint was 94% (34/36) and response to email questionnaires at other time points was 83% or greater. We found acceptable adherence (defined as greater than 70% of participants performing exercise 2 or 3 times/week) in the recommended care group with telerehabilitation but not in the recommended care group without telerehabilitation. There was a total of 24 adverse events over 108 person-months of observation. All adverse

events were mild or moderate (mainly muscle and shoulder symptoms), with the exception of one instance of elective surgery (unrelated to the person's shoulder condition).

Conclusions: Our prespecified success criteria were met or exceeded, but there was a gender imbalance toward women. It is feasible to progress to a fully powered trial, but strategies to address the gender imbalance need to be implemented.

Trial Registration: Australian New Zealand Clinical Trials Registry (ACTRN12620000248965); https://tinyurl.com/yy6eztf5

(JMIR Mhealth Uhealth 2020;8(11):e24311) doi:10.2196/24311

KEYWORDS

rotator cuff; tendinopathy; shoulder; telemedicine; telerehabilitation; randomized controlled trial; pilot; feasibility; pain; internet-delivered intervention

Introduction

Rotator cuff–related shoulder pain (RCRSP) is a common and disabling musculoskeletal condition. The estimated point prevalence of shoulder pain among adults is between 15% and 27% [1], and RCRSP is regarded as the most common cause, accounting for 70% of cases [2]. RCRSP can severely limit work and daily functions, including dressing and personal care, and can lead to substantial societal burden through utilization of health care resources and work absenteeism [1,3]. Up to 40% of those affected experience ongoing pain and disability beyond 12 months and many eventually require further interventional care (eg, injection, surgery) [1]. The estimated annual cost of managing shoulder pain is US $5,234.54 per person in 2009, and this is heavily influenced by sick leave and surgical costs [4].

Clinical practice guidelines recommend up to 12 weeks of conservative care (activity, medication, and exercise) for first presentation of RCRSP prior to considering imaging or surgery [5-7]. In contrast, high rates of imaging for first presentation of RCRSP were observed in a general practitioner (GP) database study (55%) [8] and a survey of GPs (82%) [9], both undertaken in Australia. One in 6 GPs would also refer for surgical opinion [9]. These imaging and surgical referral practices involve significant costs and, in some cases, unnecessary surgery [10], and may partly explain the doubling of rotator cuff–related surgeries and tripling of associated costs in Western Australia in just over a decade (2001-2013) [11].

Although the decision is multifactorial, one reason clinicians provide care that is not guideline recommended is to appease patients [12,13]. Pressure from patients for imaging and surgical referrals may stem from beliefs about the relevance of pathoanatomy and imaging findings [14,15]. Further, patients' beliefs about expected outcomes are strong predictors of conservative care outcomes for shoulder pain [16] and RCRSP [17]. Educating patients directly about their condition and recommended care has the potential to improve health literacy, quality of care, and health outcomes [18-20]. Additionally, patient-directed education circumvents clinician-related barriers to recommended care that can be challenging to influence.

By increasing access to guideline-recommended care, internet-based delivery of patient-directed recommended care may improve quality of care and outcomes. Internet-based delivery of patient-direct recommended care is convenient and enables care delivery to people in rural and remote regions [21-23]. However, internet delivery of health care (eg, for low back pain and chronic pain) has demonstrated heterogeneous effects on health care use and clinical outcomes [24,25]. Telerehabilitation may improve outcomes of internet-based care delivery [26] and reduce attrition [27] and may be able to replace face-to-face care that includes exercise [21] while reducing costs [28]. Further, group-based telerehabilitation may be more cost-effective and includes an opportunity for peer-to-peer support, and its outcomes are comparable with those for individual care for musculoskeletal conditions [29]. Internet-based delivery of care with or without telerehabilitation has the potential to improve quality of care and outcomes for people with RCRSP, but it is not known whether investigating these interventions in RCRSP is feasible.

The primary aim of this study was to assess the feasibility of a future substantive randomized controlled trial (RCT) comparing the effectiveness of three internet-delivered interventions for RCRSP (advice only, recommended care, and recommended care with telerehabilitation). Primary feasibility aims included assessing (1) rates of conversion and recruitment, (2) levels of adherence, (3) rate of retention, (4) incidence of adverse events, and (5) response rates to questionnaires. A secondary aim was to explore any signals of treatment effect and variability in clinical outcomes at 6 and 12 weeks.

Methods

Study Design

The study was a 3-arm, parallel-group pilot and feasibility RCT. Study design and reporting were in accordance with the Consolidated Standards of Reporting Trials (CONSORT) eHealth guidelines [30] and extension for randomized pilot and feasibility trials [31], as well as the Consensus on Exercise Reporting Template (CERT) [32]. The trial protocol was prospectively registered on February 26, 2020, at the Australian New Zealand Clinical Trials Registry (ACTRN12620000248965). Ethical approval was granted by the Monash University Human Ethics Committee (No. 22338).

Recruitment Strategy and Incentive

Thirty-six community-dwelling people with RCRSP were recruited and randomized into 3 groups (12 individuals per group). Recruitment was via a paid Facebook campaign. To compensate them for their time, participants were incentivized by receiving an Aus $100 (US $70) shopping voucher on completion of their 12-week questionnaire (regardless of other

outcome data returned). This strategy was considered important to improve participant retention [33-35].

Internet Eligibility Screening

Inclusion was based on the following question set (they were excluded if they answered one or more questions with the response indicated in parenthesis): (1) Has your shoulder problem been diagnosed by a health professional as frozen shoulder, arthritis, a labral tear, instability? (yes), (2) Is your shoulder pain a result of a shoulder dislocation? (yes), (3) Is your shoulder pain mainly around the area shown in the photos (anterolateral upper arm/shoulder pain)? (no), (4) Is your shoulder pain made worse by neck movement? (yes), (5) Is your shoulder pain brought on by moving your arm above your head? (no), and (6) Are you able to lift your arm to the height in the photo (90 degrees of elevation)? (no). The last question was designed to exclude people with massive rotator cuff tears involving multiple tendons or frozen shoulder [7]. People who were younger than 18 years, had shoulder pain for less than 3 months, had had prior surgery for their currently most symptomatic shoulder, or who had another complaint more troubling than their shoulder were excluded. People were also excluded if they indicated that they were severely depressed; taking recreational drugs, oral steroids (eg, prednisolone), or blood thinning medications (eg, warfarin); had angina, heart problems, or severe middle abdominal or upper back pain; had a history of cancer; had recent dizziness, blurred vision, slurred speech, difficulty swallowing, falls, or unsteadiness; or had a recent seizure. People excluded were advised to seek advice from their GP.

Telerehabilitation Eligibility Screening

We undertook a 20-minute teleconference (Zoom) session with participants to check answers to questions 1 to 6 described above (regardless of answers to these questions). This was only done with people who passed the remaining online screening questions.

Randomization

Participants who passed both screening stages completed an electronic consent form and were then eligible to be randomized. We utilized the two-stage Zelen randomization process in order to minimize refusal to participate related to randomization and reduce attrition (which is relatively common in the context of internet-based interventions) related to resentful demoralization in the active control group. Eligible participants were initially informed and consented to participate in a single-cohort study involving the advice-only group. The cohort was then randomized into the three groups. People randomized to one of the two intervention groups were invited to join their allocated group and were provided with information about the recommended care provided in these groups. If they agreed, they completed a second electronic consent form specific to their group allocation. Randomization was via a computer-generated random number sequence. To ensure allocation concealment, a researcher who had no contact or knowledge of any characteristics of the participants performed the randomization. Once participants were screened, the independent researcher accessed the trial database and randomized the 36 included participants. The same researcher then initiated (via emails to the participants) treatment for people in the advice-only and recommended care groups, and allocated (randomly) 3 participants to each of the 4 physiotherapists providing care in the recommended care and telerehabilitation group.

Blinding

The trial was single-blind (assessors), as all outcomes were collected via self-report questionnaires. Clinicians providing the telerehabilitation intervention were aware of treatment allocation. Because of the Zelen design, all participants perceived receiving an intervention, and participants in the recommended care groups knew their intervention contained additional components.

Baseline Assessment

Baseline assessment was via an online questionnaire and included demographic factors, medical history, site and duration of symptoms, previous treatments, and baseline outcomes (described below).

Internet Intervention Content Development

Content was based on available guidelines, systematic reviews and consensus statements [5-7,36-41]. In developing the education intervention, we also sought the views of 8 international clinical shoulder researchers (semistructured interviews) [42] and 8 patients with RCRSP (focus group).

Internet Intervention Format and Testing

Interventions included text, infographics, and videos (created with Powtoon animation software; Powtoon Ltd). Following development of the content for each of the groups, the international clinical shoulder researchers and patients were invited to provide feedback regarding the accuracy, adequacy, and clarity of the content. The Patient Education Materials Assessment Tool (PEMAT) was used to assess understandability and actionability of the multimedia education content [43]. Minor changes were made (layout and wording) based on clinician and patient feedback and the PEMAT results.

Internet-Based Interventions

Each of the internet-based interventions was 12 weeks in duration.

Advice Only

The advice-only group received education about the rotator cuff muscles and risk factors (Multimedia Appendix 1) and advice about modifying general and work-related activities (Multimedia Appendix 2) consistent with current guidelines [5,6]. Activity modification was labelled an intervention so that participants perceived that they were receiving treatment. Participants were advised to continue day-to-day and work activities if pain during these tasks was “acceptable” (a score of 4 or less on an 11-point numerical rating scale, with 10 representing worst pain imaginable) [44]. It was also acceptable for pain after activity to increase if it returned to pre-exercise levels within a reasonable period [44]. Participants were advised to stop or modify (ie, reduce volume of) activities until they could be performed with acceptable pain.

Advice With Recommended Care

Recommended care included exercise (described below) and education about the causes of RCRSP and pain mechanisms, as well as education about exercise and other treatments and their effectiveness and potential harms (Multimedia Appendix 3). Emphasis was on addressing knowledge gaps and barriers and enablers to recommended care. For example, challenging participants' understanding of the relevance of tendon structure and imaging [15], expectations from exercise interventions, and pain and exercise self-efficacy [16]. The education intervention was informed by adult learning theory [45] and evidenced-based principles of self-management and cognitive behavioral therapy [46].

Advice With Recommended Care and Telerehabilitation

The group receiving recommended care and telerehabilitation was identical to the recommended care group, but it also received a weekly telerehabilitation session with a physiotherapist via free videoconferencing software (Zoom; Zoom Video Communications Inc). Participants were provided with information about how to set up the telerehabilitation environment, including how to position their phone, tablet, or laptop (desktops were discouraged), video settings, and camera angle. Prior to starting, the telerehabilitation environment was extensively piloted for this context. In the first two rehabilitation sessions (60 minutes each), the physiotherapists presented PowerPoint slides (screensharing via Zoom) providing education about RCRSP (Session 1: Understanding RCRSP and the treatments available; Session 2: Exercise and self-management). In the remaining sessions (30 minutes each), the physiotherapists prompted discussion about pain and pathology beliefs, exercise expectations, and individual physical activity goals, as well as monitoring and providing feedback regarding exercise fidelity.

Exercise Intervention

The exercise approach was identical for the recommended care and recommended care and telerehabilitation groups and was based on available evidence and expert consensus [37-40,47]. Two exercises were included (Figure 1): shoulder elevation in standing position from 10 to 150 degrees, and external rotation in side-lying position, full range. Elevation was performed as shoulder abduction, scapular plane elevation or flexion, depending on acceptable pain response (as defined above) and patients' preference. Progression and regression were based on pain or how difficult the exercise was (Multimedia Appendix 4). If pain during exercise was 5 or greater on the 11-point numerical rating scale or the exercise was too difficult (allocated repetitions could not be completed), the following modifications were trialed (in this order): (1) reduce load if participant was using a weight, (2) reduce range of motion and/or the number of repetitions, or (3) revert to isometric hold exercise. This process was reversed to progress exercise when the pain was acceptable or the allocated volume did not achieve muscular fatigue (defined below). Participants were instructed to perform the exercises 3 times per week for 12 weeks: 3 sets of 15 repetitions and 4 seconds per cycle for isotonic exercises (2-second concentric and 2-second eccentric phase) or 6 sets of 30 seconds for isometric exercise, with a 2-minute rest between each set. Fatigue was defined as self-reported inability to complete a further complete repetition. Participants were encouraged to use a heavier weight if they felt they could do 17 repetitions or more, or a lighter weight if they could do 13 repetitions or fewer). Participants were advised to adjust the weight in 0.5 kg or 1 kg increments and use their own dumbbells at home or an empty container (eg, large plastic milk container) filled with water until it approximated the desired weight. Participants entered pain and fatigue data onto the website (Multimedia Appendix 5) prior to commencing the exercise session and after each set of exercise, and the algorithm provided guidance on progression and regression.

Figure 1. External rotation and elevation exercises.

Behavioral Strategies for Intervention Groups

Strategies to improve adherence included an interactive website-based exercise progression and regression tool with built-in electronic exercise diary (Multimedia Appendix 5), allowing patients some choice in exercise selection (ie, elevation plane), challenging cognitions and beliefs that might impact adherence (eg, pain beliefs or fear avoidance, as described above), and providing regular feedback and monitoring (telerehabilitation group) [48].

Physiotherapist Recruitment and Training

Four registered physiotherapists working in primary care settings with experience managing shoulder conditions were recruited to deliver the telerehabilitation interventions. The physiotherapists participated in four 1-hour training sessions via teleconference to learn about the intervention and standardize the delivery of care.

Primary Feasibility Outcomes

The following outcomes (based on a high-quality trial undertaken in the same geographical region and using similar recruitment strategies [49]) were used to determine feasibility (ie, yes, no, or modifications required) for a substantive RCT:

- the number of eligible participants who consented and were randomized (feasible defined as 20% or greater).
- exercise adherence—participants were asked whether they completed 0, 1, 2, or 3 or more of the 3 prescribed sessions per week; acceptable adherence was defined as 70% or more participants in a group completing 2 or more prescribed exercise sessions) per week.
- rate of retention (feasible defined as 80% or greater).
- response rates to questionnaire outcomes (feasible defined as 80% or greater).
- incidence of adverse events—adverse events were defined as any symptom experienced during the trial that may or may not be causally related to the intervention; the frequency (number of participants and number of adverse events), type (eg, muscle soreness), and severity (mild: lasting less than 7 consecutive days; moderate: lasting 7 consecutive days or longer; severe: results in death, life-threatening complication, hospitalization, surgery, permanent or temporary physical disability, congenital abnormalities, or any findings the research team believed could lead to significant health hazards).

Adverse events and adherence were reported via email questionnaire at 4, 8, and 12 weeks.

Outcomes for Economic Evaluation

Outcomes for economic evaluation were included to assess the feasibility of collection in a future full-scale trial (feasible defined as a response rate of 80% or greater). Health care use related to RCRSP was assessed with a patient questionnaire. Productivity (including absenteeism and presenteeism) was measured using the iMTA Productivity Cost Questionnaire [50]. These outcomes were collected via email questionnaire at 4, 8, and 12 weeks. Costs were divided into direct costs (intervention related and other) and indirect costs (including absenteeism and presenteeism). Physiotherapists' hourly rate was assumed to be Aus $150 (US $106) and average Australian hourly rate to be Aus $60 (US $42).

Secondary Clinical Outcomes

Clinical outcomes were assessed at baseline, 6 weeks, and 12 weeks via email questionnaire as follows:

- pain and function—measured with the Shoulder Pain and Disability Index (SPADI), a validated questionnaire [51] that has been used extensively among people with RCRSP; the minimal clinically important difference (MCID) for SPADI among people with RCRSP is reported to be between 8 and 13 points [52].
- global rating of change—measured using the patient Global Rating of Change (GROC) 11-point Likert scale [53]; participants were asked to rate how their shoulder pain had changed since receiving the intervention.
- worst pain in the last 7 days—measured using the 100-mm visual analog scale (VAS; zero=no pain, 100=worst pain possible).
- health-related quality of life—measured with the EuroQol 5D-5L (EQ5D), a validated and reliable tool [54], including 5 domains (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression).
- kinesiophobia (fear of movement)—measured with the Tampa Scale for Kinesiophobia (TSK), which has been validated among people with musculoskeletal pain [55].
- pain catastrophizing—measured with the Pain Catastrophizing Scale (PCS), a validated measure of pain catastrophizing [56].
- pain self-efficacy—measured with the reliable, valid, and responsive short form of the Pain Self-Efficacy Questionnaire (PSEQ) [57].
- RCRSP knowledge—assessed with a custom knowledge test at 0 (pre- and immediately post-education), 6, and 12 weeks. The questionnaire was developed based on the patient knowledge questionnaire for patients with osteoarthritis (OA) [58] and included 16 multiple-choice questions (6 questions about the disease, eg, risks and symptoms, and 10 questions about treatments, and treatment efficacies and harms). Psychometric properties of this scale were evaluated and deemed acceptable.

If participants did not complete email questionnaires within 7 days, they were reminded via email (2 emails/week) and later via a phone call.

Statistical Analysis

Data were analyzed using SPSS (version 25; IBM Corp). Frequencies and proportions were found for categorical data, means and SDs for continuous data, and medians and IQRs for ordinal data. This pilot RCT was not powered to detect comparative treatment effects; however, within-group mean differences and 95% CIs as well as within-group standardized mean differences (SMDs, mean difference/pooled SD) were reported for change between baseline and 6 weeks and between baseline and 12 weeks. The SMD was interpreted in the following way: 0.2=small effect, 0.5=moderate effect, 0.8=large effect, and 1.2=very large effect [59]. The GROC was dichotomized, where a successful outcome was defined as "moderately better," "much better," or "very much better" on this scale [53].

Results

Demographic data for the 36 participants recruited are shown in Table 1. The mean age of participants was 53.9 (SD 12.0) years and 89% (32/36) of the cohort was female.

Table 1. Demographic information of study participants.

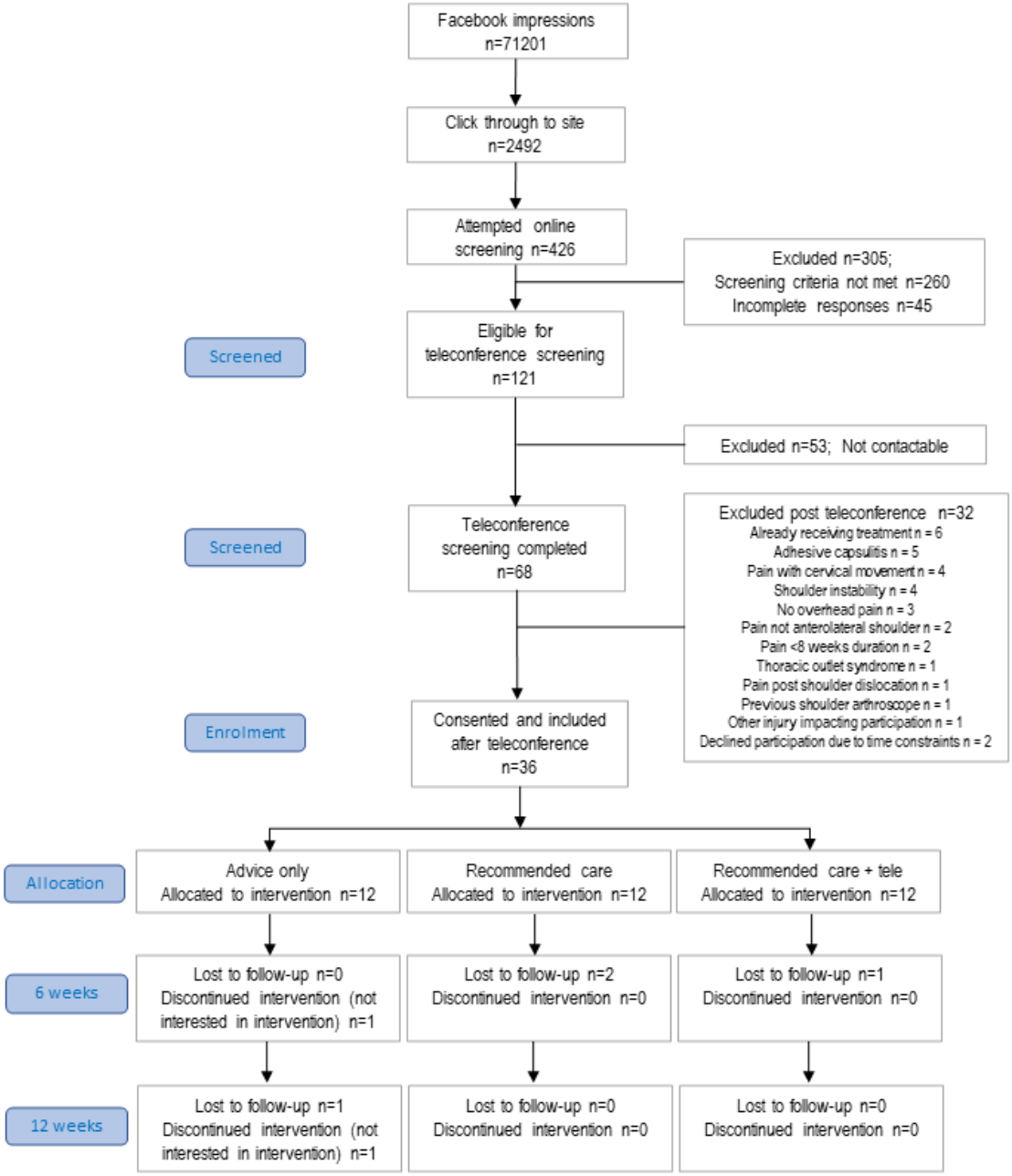
Demographic factor	Advice only (n=12)	Recommended care (n=12)	Recommended care and telerehabilitation (n=12)
Age (years), mean (SD)	53.7 (11.5)	51.3 (13.7)	56.6 (11.0)
Female, n (%)	11 (92)	10 (83)	11 (92)
Height (cm), mean (SD)	166.3 (7.0)	170.8 (11.8)	165.1 (8.2)
Mass (kg), mean (SD)	84.3 (24.7)	74.9 (16.9)	76.4 (12.6)
BMI (kg/m^2), mean (SD)	30.9 (10.3)	25.9 (5.7)	28.0 (4.2)
Employed, n (%)	6 (50)	8 (67)	5 (42)
Residence, n (%)			
Urban	7 (58)	9 (75)	9 (75)
Other urban	4 (25)	2 (17)	3 (25)
Rural	1 (8)	1 (8)	0 (0)
Affected/worst side is right side, n (%)	8 (67)	6 (50)	7 (58)
Duration of symptoms (weeks), mean (SD)	27.6 (17.1)	42.5 (17.7)	33.8 (20.0)
Prior exercise treatment, n (%)	8 (67)	8 (67)	9 (75)
Prior activity modification, n (%)	2 (17)	5 (42)	6 (50)
Prior shoulder imaging, n (%)	3 (25)	6 (50)	3 (25)
SPADI, mean (SD)	30.6 (17.7)	37.3 (16.7)	41.8 (19.0)
Worst pain in the previous week, mean (SD)	5.4 (2.8)	7.5 (2.0)	6.3 (3.8)
Comorbidities, n (%)			
Osteoarthritis	5 (42)	0 (0)	3 (25)
Rheumatoid arthritis	0 (0)	2 (17)	0 (0)
Inflammatory bowel disease	1 (8)	0 (0)	0 (0)
Fibromyalgia	0 (0)	2 (17)	0 (0)
Hypertension	1 (8)	0 (0)	1 (8)
Hypercholesterolemia	0 (0)	2 (17)	5 (42)
Diabetes	1 (8)	1 (8)	3 (25)

Primary Feasibility Outcomes

Rate of Conversion and Recruitment

The CONSORT flow diagram is shown in Figure 2. The total cost of the Facebook advertising campaign was Aus $2,146.02 (US $1512.94). The campaign was active for 21 days over 4 weeks—average cost of Aus $59.61 (US $42.03) per participant recruited)—which generated 71,201 impressions and 2492 clicks to the trial site. There were 68 potentially eligible participants who were screened via teleconference, of which 38 were eligible and 36 consented to participate (rate of conversion=36/38=95% of eligible participants recruited). All 36 participants were recruited in 1 calendar month.

Figure 2. Consolidated Standards of Reporting Trials (CONSORT) flow diagram. tele: telerehabilitation.



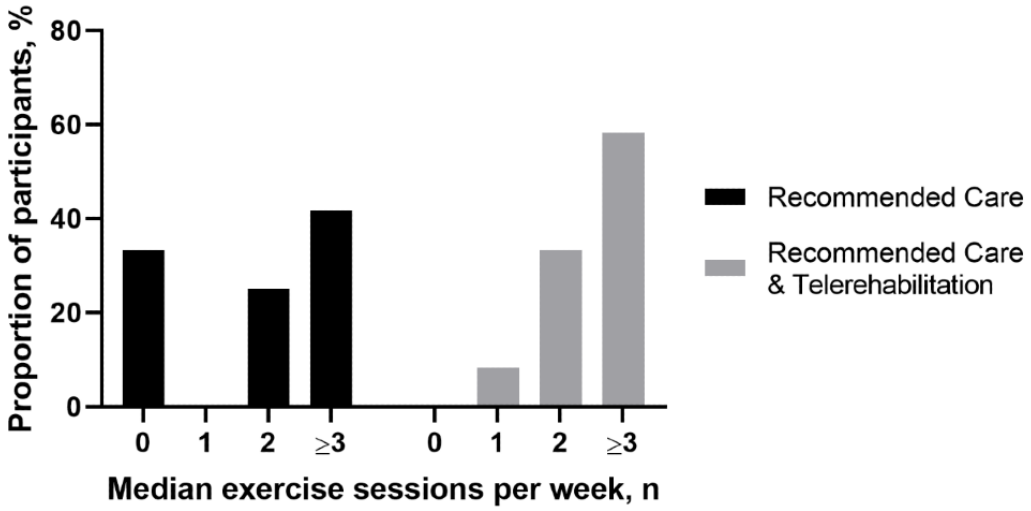
Retention and Email Questionnaire Response Rates

The questionnaire data at 12 weeks were submitted via email by 34 participants (34/36, 94% retention). Response to email questionnaires at 4 weeks (32/36, 88%), 6 weeks (32/36, 88%), and 8 weeks (30/36, 83%) were also acceptable.

Exercise and Telerehabilitation Adherence

Acceptable exercise adherence was achieved in the recommended care group (11/12, 92%) but not without telerehabilitation (8/12, 67%) (Figure 3). One-third (4/12, 33%) of people in the recommended care without telerehabilitation group performed no exercise at all. A mean of 10.6 (SD 1.4) out of 12 (88%) teleconference sessions were attended.

Figure 3. Exercise adherence in patients receiving recommended care and recommended care with telerehabilitation.



Adverse Events

Adverse events reported were classified as either mild or moderate (Table 2). There was 1 serious adverse event—a participant in the recommended care with telerehabilitation group had elective surgery (brain surgery unrelated to RCRSP) and was absent from the trial for approximately 3 weeks. Other adverse events in all 3 groups could be categorized as either increased shoulder symptoms or muscle soreness around the shoulder, with the exception of 2 incidences of back pain in the advice-only group.

Co-interventions

Two people (2/12, 17%) in each group reported between 14 and 16 incidences of co-interventions during the trial week (Table 2). Therapy sessions were the most common type of co-interventions, including physiotherapy (only utilized by people in groups that did not include telerehabilitation), chiropractic, massage therapy, or GP visit. Number of days of medication use ranged from 14 to 27 equating to 1.3% to 2.5% of person-days (12 participants x 90-day study period = 1080 person-days).

Table 2. Adverse events and co-interventions.

Adverse event or co-intervention measure	Advice only	Recommended care	Recommended care and telerehabilitation
Adverse events, n (%)			
Participants reporting adverse event	6 (50)	4 (33)	6 (50)
Total number of adverse events	10	7	8
Mild	7 (70)	4 (57)	6 (75)
Moderate	3 (30)	3 (43)	1 (12)
Serious	0 (0)	0 (0)	1 (12)
Co-intervention, n (%)			
Participants using co-interventions	2 (17)	2 (17)	2 (17)
Total number of co-interventions	14	16	15
Therapy sessions	9 (64)	9 (56)	9 (60)
Other interventions	1 (7)	5 (31)	3 (20)
Episodes of medication use	4 (29)	2 (13)	3 (20)
Participants using medication, n (%)	2 (17)	2 (17)	2 (17)
Total number of days of medication use[a]	27	14	24

[a]From a total of 1080 person-days (12 participants x 90-day study period).

Feasibility of Future Economic Evaluation

The response to health service use questionnaires and iMTA Productivity Cost was acceptable (greater than 80%), indicating it is just as feasible to collect economic outcomes from this population as it is clinical outcomes. Direct costs were Aus $195.64 (SD 224.25; US $137.93), absenteeism was Aus $223.33 (SD 864.41; US $157.45), and presenteeism was Aus $430.23 (SD 752.61; US $303.31) per person across the cohort.

Secondary clinical outcomes

Data for secondary outcomes are presented in Figure 4 and Table 3.

The median knowledge test score was between 6 and 9 of a possible 16 points on the baseline test (Figure 5). Immediately after exposure to the education intervention, the median increased to 14 or 15 points in the groups receiving the recommended care, with a smaller increase in the advice-only group.

Figure 4. Global rating of change.

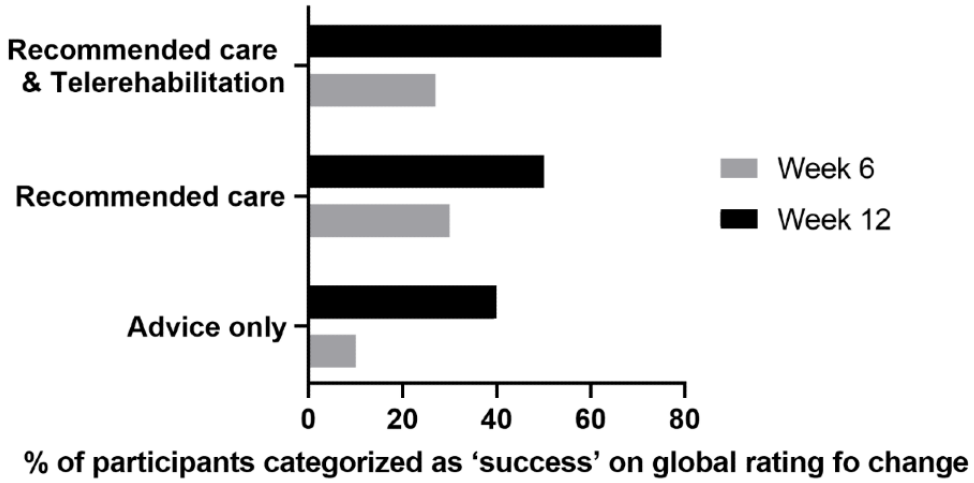


Table 3. Secondary outcomes (reported as mean [SD]).

Outcome	Advice only	Recommended care	Recommended care and telerehabilitation
SPADI[a]			
Baseline	30.6 (17.7)	37.3 (16.7)	41.8 (19.1)
Week 6	32.9 (19.9)	27.3 (18.5)	15.0 (9.5)
Week 12	26.6 (22.3)	21.7 (17.8)	12.9 (6.9)
Mean difference at 6 weeks (95% CI)	2.3 (–9.5 to 14.1)	–14.0 (–26.7 to –1.3)	–26.9 (–39.2 to –14.5)
Mean difference at 12 weeks (95% CI)	–4.8 (–20.3 to 10.8)	–16.0 (–26.0 to –6.0)	–28.9 (–40.9 to –28.7)
SMD[b] baseline-week 6	0.1	0.8	1.4
SMD baseline-week 12	0.3	0.9	1.4
Worst pain in previous 7 days			
Baseline	56.8 (17.9)	51.6 (22.4)	59.7 (21.1)
Week 6	55.7 (22.2)	41.5 (22.2)	31.9 (23.1)
Week 12	41.8 (23.1)	44.8 (28.1)	28.1 (25.6)
Mean difference at 6 weeks (95% CI)	–0.2 (–11.3 to 10.9)	–9.3 (–34.9 to 16.29)	–26.8 (–44.8 to –8.9)
Mean difference at 12 weeks, (95% CI)	–15.8 (–33.1 to 1.5)	–3.2 (–19.7 to 13.4)	–31.6 (–49.89 to –13.28)
SMD baseline-week 6	0.1	0.6	1.0
SMD baseline-week 12	0.2	0.6	1.0
TSK[c]			
Baseline	36.0 (7.0)	35.2 (5.8)	35.5 (6.7)
Week 6	37.3 (6.6)	31.5 (8.9)	32.6 (7.5)
Week 12	36.3 (6.5)	30.6 (5.7)	31.1 (6.6)
Mean difference at 6 weeks (95% CI)	–0.2 (–4.2 to 3.8)	–4.8 (–8.8 to –0.8)	–3.6 (–7.2 to 0.1)
Mean difference at 12 weeks, (95% CI)	–1.9 (–5.9 to 2.1)	–5.0 (–7.0 to –3.0)	–4.4 (–8.8 to –0.1)
SMD baseline-week 6	0.0	0.6	0.5
SMD baseline-week 12	0.3	0.8	0.6
PCS[d]			
Baseline	7.3 (9.2)	4.6 (5.8)	4.8 (4.1)
Week 6	8.8 (10.5)	6.0 (10.5)	3.7 (3.3)
Week 12	5.1 (0.5 to 9.7)	2.7 (1.1 to 4.2)	4.0 (1.0 to 7.0)
Mean difference at 6 weeks (95% CI)	0.6 (–1.1 to 2.3)	0.1 (–1.8 to 2.0)	–1.6 (–3.1 to 0.0)
Mean difference at 12 weeks (95% CI)	–3.4 (–7.3 to 0.5)	–2.0 (–4.6 to 0.6)	–0.8 (–3.3 to 1.6)
SMD baseline-week 6	0.0	0.0	0.4
SMD baseline-week 12	0.4	0.5	0.2
PSEQ[e]			
Baseline	50.4 (9.4)	50.1 (9.3)	52.3 (6.9)
Week 6	49.7 (6.4)	51 (7.6)	55.1 (4.1)
Week 12	50.5 (7.9)	54.8 (6.5)	55.6 (5.8)
Mean difference at 6 weeks (95% CI)	–1.1 (–5.1 to 2.9)	1.7 (–7.2 to 10.5)	3.6 (0.8 to 6.3)
Mean difference at 12 weeks, (95% CI)	1.3 (–4.1 to 6.7)	4.2 (–1.6 to 9.9)	3.3 (1.0 to 5.6)
SMD baseline-week 6	0.1	0.2	0.6
SMD baseline-week 12	0.2	0.5	0.5
EQ5D[f]			

Outcome	Advice only	Recommended care	Recommended care and telerehabilitation
Baseline	0.76 (0.11)	0.74 (0.13)	0.74 (0.12)
Week 6	0.73 (0.12)	0.74 (0.12)	0.77 (0.11)
Week 12	0.73 (0.09)	0.77 (0.13)	0.78 (0.07)
Mean difference at 6 weeks (95% CI)	–0.02 (–0.10 to 0.06)	0.02 (–0.08 to 0.12)	0.05 (–0.02 to 0.12)
Mean difference at 12–weeks, (95% CI)	–0.02 (–0.14 to 0.10)	0.01 (–0.10 to 0.12)	0.03 (–0.02 to 0.09)
SMD baseline-week 6	0.2	0.2	0.4
SMD baseline-week 12	0.2	0.1	0.3

[a]SPADI: shoulder pain and disability index.

[b]SMD: standardized mean difference.

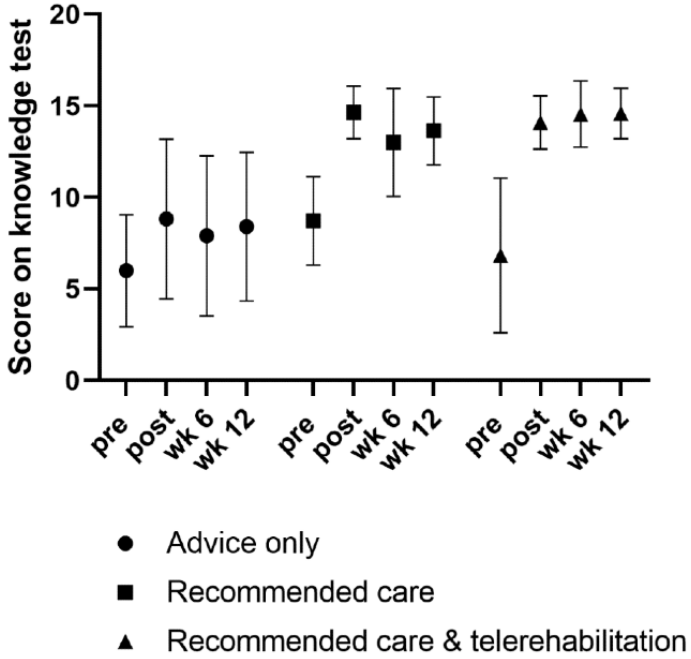
[c]TSK: Tampa Scale for Kinesiophobia.

[d]PCS: Pain Catastrophizing Scale.

[e]PSEQ: Pain Self-Efficacy Questionnaire.

[f]EQ5D: EuroQol 5D-5L.

Figure 5. Knowledge test change over time. post: postintervention; pre: preintervention; wk: week.



Discussion

Principal Findings

Based on our predefined progression criteria, conducting a full-scale RCT to evaluate delivery of care via the internet and telerehabilitation for RCRSP is feasible. Prespecified rates of conversion (36/38, 96%; criterion of greater than 20%), recruitment (36 per month), and retention (34/36, 94%; criterion of greater than 80%) were high, and there were acceptable responses to questionnaire outcomes at all time points (30/36, 83% or higher; criterion of greater than 80%). Adherence was acceptable (2 to 3 sessions per week among more than 70% of participants) for the group receiving recommended care with telerehabilitation (11/12, 92%) but not for the group receiving recommended care without telerehabilitation (8/12, 67%). The trial was not powered to investigate between-group differences, but we have reported descriptive variability and within-group change data for clinical outcomes in accordance with pilot trial recommendations [60].

Although recruitment was efficient using our social media strategy, there was a gender imbalance favoring women (32/36, 89%). A systematic review of health behavior change interventions also reported much higher participation among women even though both genders were targeted via a social media campaign [61]. Addressing gender balance is an important consideration for a future full-scale trial to ensure that the findings are generalizable to the population with RCRSP. There are reports that men are more likely to respond to Facebook ads that utilize concise text and appeal to leadership themes and masculine themes, and less likely to respond to unisex ads [62]. Facebook also enables targeting ads only to men and this could be used during a full-scale trial if gender imbalance emerged.

By removing cost and travel barriers, our internet and remote interventions may have promoted health-seeking behavior among people who would not ordinarily seek care. Our participants were slightly younger (mean age 54 years versus 60 years) and had slightly lower SPADI scores (37 versus 43) compared with a large, high-quality trial that recruited people with RCRSP via primary care from the same geographic region [49]. However, more people in our trial had tried exercise or physiotherapy treatment (69% versus 38%) [49], indicating that most of our cohort is representative of people who seek care for RCRSP. Nevertheless, strategies to increase the RCRSP severity of people recruited (eg, a minimum SPADI score as an inclusion criterion) may increase the extent to which our cohort is representative of people who seek care for this condition in primary care.

The Zelen randomization design that we used (theoretically) provides two key design advantages: (1) it improves recruitment rates by avoiding consent refusal related to simple randomization, and (2) it reduces attrition (and therefore selection bias) related to resentful demoralization among people in the control group [63]. This approach may have contributed to the efficiency of recruitment and high retention in this feasibility trial. Opting to remain in the original group rather than joining the group that people are randomized to (also referred to as crossover) is another potential issue with the Zelen design [63], but 100% of people accepted the offer of randomization in our trial. A potential limitation of the Zelen design is that people offered additional treatment perceive this as superior to the original care they consented to, but this can be mitigated by careful wording of the trial information to introduce uncertainty about the comparative efficacy of the interventions being tested. Overall, we recommend that researchers consider Zelen's design for internet-based interventions where recruitment and retention may be challenging (eg, where there is no human contact).

Efficient recruitment and high retention rates may also be related to the monetary incentive (Aus $100 [US $70] for 12-week outcomes). Previous systematic reviews have reported that monetary incentives can improve retention rates for studies that utilise postal or electronic questionnaires [33-35]. We employed this strategy because of the substantial attrition rates that have been reported in some trials investigating internet interventions aimed at improving health outcomes, including interventions for musculoskeletal conditions [26,64,65]. For example, 45% attrition was reported at 3 months for an internet-delivered physical activity intervention for knee OA [66]. Although we have no knowledge of rates of attrition without the monetary incentive, attrition may be an issue for our interventions, particularly the advice-only intervention (active control).

Contact with a physiotherapist may explain the higher exercise adherence in the group receiving telerehabilitation. Consistent with our findings, there are reports of low adherence for internet-only interventions [67], and a recent study found that website-based exercise with remote physiotherapist support improved adherence (at 4 weeks) compared with the provision of a leaflet explaining the exercises [68]. Participants in our study may have perceived value in telerehabilitation (eg, improved explanation, reassurance, feedback, monitoring, individualization, and peer support). Alternatively, our findings may be explained by Hawthorne effects [69]. We used behavioral strategies to improve adherence in the internet intervention (eg, electronic exercise diary, addressing beliefs), although "persuasive technologies" such as electronic reminders or tailoring to individual users' needs or personalities may further improve adherence in the group receiving recommended care without telerehabilitation [70].

Comparison With Prior Work

There are numerous internet-only delivery of care interventions targeted toward people with chronic pain as well as OA, but feasibility (retention and adherence) and outcomes have been mixed [26,66,71]. There is also evidence that care for people with musculoskeletal conditions can be successfully delivered via telerehabilitation. Cottrell et al [21] reviewed trials investigating telerehabilitation interventions for musculoskeletal conditions (knee OA, neck pain) and after shoulder and knee joint surgery. The review concluded that the telerehabilitation-only interventions, which are comparable with our recommended care and telerehabilitation group, were equivalent to face-to-face care for function outcomes. Our intervention is novel in that it blends internet, telerehabilitation, and group-based care delivery for RCRSP.

Strengths

The content of our education intervention was informed by stakeholders, including patients, clinicians, or researchers with expertise in shoulder pain management [42]. Stakeholders helped to determine the education needs for people with RCRSP and subsequently assessed accuracy and clarity of the content and delivery modes. The exercises were based on consensus statements [37,38], systematic reviews [36,39-41], practice guidelines [5-7], and a protocol for a high-quality trial currently underway in the United Kingdom [72]. The intervention incorporated only two exercises to ensure simplicity for the internet-only delivery mode and included behavioral strategies to increase exercise adherence [72].

Limitations

This trial design has limitations. First, generalizability to a primary care population may be improved by implementing strategies to improve the gender balance and recruit people with greater RCRSP severity (as discussed above). Second, we did not measure adherence to recommendations to modify activity behavior (the only intervention component in the advice-only group), which may help to explain findings in this group. Third, the findings may be influenced by placebo and contextual factors given that participants in the intervention groups were aware that they were allocated to an intervention involving additional evidence-based care. In our pragmatic trial design, we chose a control intervention that represents an acceptable standard of care [73]. Fourth, diagnosis was based on remote screening (online and teleconference). Although it was based on guidelines [7], this remote method has not been tested.

Conclusion

Our prespecified success criteria were met or exceeded but there was a gender imbalance toward women. It is feasible to progress

to a fully powered trial, but strategies to address the gender imbalance need to be implemented.

Acknowledgments

Funding for this study was provided by an Arthritis Australia Project Grant (Aus $30,000 [US $21,150) and an Australian Physiotherapy Association Physiotherapy Research Fund Seed Grant (Aus $11,840 [US $8347]).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Educational materials for advice-only group.
[PDF File (Adobe PDF File), 835 KB - mhealth_v8i11e24311_app1.pdf]

Multimedia Appendix 2
Activity-modification guide for advice-only group.
[PDF File (Adobe PDF File), 983 KB - mhealth_v8i11e24311_app2.pdf]

Multimedia Appendix 3
Educational materials for recommended care group.
[PDF File (Adobe PDF File), 976 KB - mhealth_v8i11e24311_app3.pdf]

Multimedia Appendix 4
Progression and regression criteria for exercise.
[PDF File (Adobe PDF File), 62 KB - mhealth_v8i11e24311_app4.pdf]

Multimedia Appendix 5
Website interface of exercise progression and regression tool with built-in electronic exercise diary.
[PNG File , 309 KB - mhealth_v8i11e24311_app5.png]

References

1. Ottenheijm RP, Joore MA, Walenkamp GH, Weijers RE, Winkens B, Cals JW, et al. The Maastricht Ultrasound Shoulder pain trial (MUST): Ultrasound imaging as a diagnostic triage tool to improve management of patients with non-chronic shoulder pain in primary care. BMC Musculoskelet Disord 2011 Jul 8;12(1). [doi: 10.1186/1471-2474-12-154]
2. van der Windt DA, Koes BW, Boeke AJ, Devillé W, De Jong BA, Bouter LM. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996 Sep;46(410):519-523 [FREE Full text] [Medline: 8917870]
3. Lewis J. Rotator cuff related shoulder pain: Assessment, management and uncertainties. Manual Therapy 2016 Jun;23:57-68. [doi: 10.1016/j.math.2016.03.009]
4. Virta L, Joranger P, Brox JI, Eriksson R. Costs of shoulder pain and resource use in primary health care: a cost-of-illness study in Sweden. BMC Musculoskelet Disord 2012 Feb 10;13(1):17 [FREE Full text] [doi: 10.1186/1471-2474-13-17] [Medline: 22325050]
5. Diercks R, Bron C, Dorrestijn O, Meskers C, Naber R, de Ruiter T, et al. Guideline for diagnosis and treatment of subacromial pain syndrome. Acta Orthopaedica 2014 May 21;85(3):314-322. [doi: 10.3109/17453674.2014.920991]
6. Hopman K. Clinical practice guidelines for the management of rotator cuff syndrome in the workplace. Port Macquarie (Australia): University of New South Wales 2013:80.
7. Kulkarni R, Gibson J, Brownson P, Thomas M, Rangan A, Carr AJ, et al. Subacromial shoulder pain. Shoulder & Elbow 2015 Mar 31;7(2):135-143. [doi: 10.1177/1758573215576456]
8. Naunton J, Harrison C, Britt H, Haines T, Malliaras P. General practice management of rotator cuff related shoulder pain: A reliance on ultrasound and injection guided care. PLoS ONE 2020 Jan 13;15(1):e0227688. [doi: 10.1371/journal.pone.0227688]
9. Buchbinder R, Staples MP, Shanahan EM, Roos JF. General Practitioner Management of Shoulder Pain in Comparison with Rheumatologist Expectation of Care and Best Evidence: An Australian National Survey. PLoS ONE 2013 Apr 16;8(4):e61243. [doi: 10.1371/journal.pone.0061243]
10. Traeger AC, Moynihan RN, Maher CG. Wise choices: making physiotherapy care more valuable. Journal of Physiotherapy 2017 Apr;63(2):63-65. [doi: 10.1016/j.jphys.2017.02.003]

11. Thorpe A, Hurworth M, O'Sullivan P, Mitchell T, Smith A. Rising trends in surgery for rotator cuff disease in Western Australia. ANZ J Surg 2016 Aug 04;86(10):801-804. [doi: 10.1111/ans.13691]
12. Egerton T, Diamond L, Buchbinder R, Bennell K, Slade SC. Barriers and enablers in primary care clinicians' management of osteoarthritis: protocol for a systematic review and qualitative evidence synthesis. BMJ Open 2016 May 27;6(5):e011618. [doi: 10.1136/bmjopen-2016-011618]
13. Slade SC, Kent P, Bucknall T, Molloy E, Patel S, Buchbinder R. Barriers to primary care clinician adherence to clinical guidelines for the management of low back pain: protocol of a systematic review and meta-synthesis of qualitative studies. BMJ Open 2015 Apr 21;5(4):e007265-e007265. [doi: 10.1136/bmjopen-2014-007265]
14. Cuff A, Littlewood C. Subacromial impingement syndrome – What does this mean to and for the patient? A qualitative study. Musculoskeletal Science and Practice 2018 Feb;33:24-28. [doi: 10.1016/j.msksp.2017.10.008]
15. Gillespie MA, Mącznik A, Wassinger CA, Sole G. Rotator cuff-related pain: Patients' understanding and experiences. Musculoskeletal Science and Practice 2017 Aug;30:64-71. [doi: 10.1016/j.msksp.2017.05.009]
16. Chester R, Jerosch-Herold C, Lewis J, Shepstone L. Psychological factors are associated with the outcome of physiotherapy for people with shoulder pain: a multicentre longitudinal cohort study. Br J Sports Med 2016 Jul 21;52(4):269-275. [doi: 10.1136/bjsports-2016-096084]
17. Kuhn JE, Dunn WR, Sanders R, An Q, Baumgarten KM, Bishop JY, et al. Effectiveness of physical therapy in treating atraumatic full-thickness rotator cuff tears: a multicenter prospective cohort study. Journal of Shoulder and Elbow Surgery 2013 Oct;22(10):1371-1379. [doi: 10.1016/j.jse.2013.01.026]
18. Dwyer DS, Liu H. The impact of consumer health information on the demand for health services. The Quarterly Review of Economics and Finance 2013 Feb;53(1):1-11. [doi: 10.1016/j.qref.2012.11.002]
19. Louw A, Diener I, Butler DS, Puentedura EJ. The Effect of Neuroscience Education on Pain, Disability, Anxiety, and Stress in Chronic Musculoskeletal Pain. Archives of Physical Medicine and Rehabilitation 2011 Dec;92(12):2041-2056. [doi: 10.1016/j.apmr.2011.07.198]
20. Smeets RJ, Vlaeyen JW, Kester AD, Knottnerus JA. Reduction of Pain Catastrophizing Mediates the Outcome of Both Physical and Cognitive-Behavioral Treatment in Chronic Low Back Pain. The Journal of Pain 2006 Apr;7(4):261-271. [doi: 10.1016/j.jpain.2005.10.011]
21. Cottrell MA, Galea OA, O'Leary SP, Hill AJ, Russell TG. Real-time telerehabilitation for the treatment of musculoskeletal conditions is effective and comparable to standard practice: a systematic review and meta-analysis. Clin Rehabil 2016 May 02;31(5):625-638. [doi: 10.1177/0269215516645148]
22. Cottrell MA, Russell TG. Telehealth for musculoskeletal physiotherapy. Musculoskeletal Science and Practice 2020 Aug;48:102193. [doi: 10.1016/j.msksp.2020.102193]
23. Dekkers T, Melles M, Groeneveld BS, de Ridder H. Web-Based Patient Education in Orthopedics: Systematic Review. J Med Internet Res 2018 Apr 23;20(4):e143. [doi: 10.2196/jmir.9013]
24. Macea DD, Gajos K, Daglia Calil YA, Fregni F. The Efficacy of Web-Based Cognitive Behavioral Interventions for Chronic Pain: A Systematic Review and Meta-Analysis. The Journal of Pain 2010 Oct;11(10):917-929. [doi: 10.1016/j.jpain.2010.06.005]
25. Li Y, Tse MYM. An Online Pain Education Program for Working Adults: Pilot Randomized Controlled Trial. J Med Internet Res 2020 Jan 14;22(1):e15071. [doi: 10.2196/15071]
26. Bender JL, Radhakrishnan A, Diorio C, Englesakis M, Jadad AR. Can pain be managed through the Internet? A systematic review of randomized controlled trials. Pain 2011;152(8):1740-1750. [doi: 10.1016/j.pain.2011.02.012]
27. Andersson G. Using the Internet to provide cognitive behaviour therapy. Behaviour Research and Therapy 2009 Mar;47(3):175-180. [doi: 10.1016/j.brat.2009.01.010]
28. Fatoye F, Gebrye T, Fatoye C, Mbada CE, Olaoye MI, Odole AC, et al. The Clinical and Cost-Effectiveness of Telerehabilitation for People With Nonspecific Chronic Low Back Pain: Randomized Controlled Trial. JMIR Mhealth Uhealth 2020 Jun 24;8(6):e15375. [doi: 10.2196/15375]
29. O'Keeffe M, Hayes A, McCreesh K, Purtill H, O'Sullivan K. Are group-based and individual physiotherapy exercise programmes equally effective for musculoskeletal conditions? A systematic review and meta-analysis. Br J Sports Med 2016 Jun 24;51(2):126-132. [doi: 10.1136/bjsports-2015-095410]
30. Eysenbach G. CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. J Med Internet Res 2011 Dec 31;13(4):e126. [doi: 10.2196/jmir.1923]
31. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24:i5239. [doi: 10.1136/bmj.i5239]
32. Slade SC, Dionne CE, Underwood M, Buchbinder R. Consensus on Exercise Reporting Template (CERT): Explanation and Elaboration Statement. Br J Sports Med 2016 Oct 05;50(23):1428-1437. [doi: 10.1136/bjsports-2016-096651]
33. Bower P, Brueton V, Gamble C, Treweek S, Smith CT, Young B, et al. Interventions to improve recruitment and retention in clinical trials: a survey and workshop to assess current practice and future priorities. Trials 2014 Oct 16;15:399 [FREE Full text] [doi: 10.1186/1745-6215-15-399] [Medline: 25322807]

34. Edwards P, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2009 Jul 08(3):MR000008. [doi: 10.1002/14651858.MR000008.pub4] [Medline: 19588449]
35. Brueton VC, Tierney JF, Stenning S, Meredith S, Harding S, Nazareth I, et al. Strategies to improve retention in randomised trials: a Cochrane systematic review and meta-analysis. BMJ Open 2014 Feb 04;4(2):e003821. [doi: 10.1136/bmjopen-2013-003821]
36. Haik MN, Alburquerque-Sendín F, Moreira RFC, Pires ED, Camargo PR. Effectiveness of physical therapy treatment of clearly defined subacromial pain: a systematic review of randomised controlled trials. Br J Sports Med 2016 Jun 10;50(18):1124-1134. [doi: 10.1136/bjsports-2015-095771]
37. Klintberg IH, Cools AMJ, Holmgren TM, Holzhausen AG, Johansson K, Maenhout AG, et al. Consensus for physiotherapy for shoulder pain. International Orthopaedics (SICOT) 2014 Dec 31;39(4):715-720. [doi: 10.1007/s00264-014-2639-9]
38. Littlewood C, Bateman M, Connor C, Gibson J, Horsley I, Jaggi A, et al. Physiotherapists' recommendations for examination and treatment of rotator cuff related shoulder pain: A consensus exercise. PPR 2019 Sep 13;40(2):87-94. [doi: 10.3233/ppr-190129]
39. Malliaras P, Johnston R, Street G, Littlewood C, Bennell K, Haines T, et al. The Efficacy of Higher Versus Lower Dose Exercise in Rotator Cuff Tendinopathy: A Systematic Review of Randomized Controlled Trials. Archives of Physical Medicine and Rehabilitation 2020 Oct;101(10):1822-1834. [doi: 10.1016/j.apmr.2020.06.013]
40. Naunton J, Street G, Littlewood C, Haines T, Malliaras P. Effectiveness of progressive and resisted and non-progressive or non-resisted exercise in rotator cuff related shoulder pain: a systematic review and meta-analysis of randomized controlled trials. Clin Rehabil 2020 Jun 22;34(9):1198-1216. [doi: 10.1177/0269215520934147]
41. Page M, Green S, McBain B, Surace SJ, Deitch J, Lyttle N, et al. Manual therapy and exercise for rotator cuff disease. Cochrane Database Syst Rev 2016 Jun 10(6):CD012224. [doi: 10.1002/14651858.CD012224] [Medline: 27283590]
42. White J, Auliffe SM, Jepson M, Burstein F, Hopman R, Morrissey D, et al. 'There is a very distinct need for education' among people with rotator cuff tendinopathy: An exploration of health professionals' attitudes. Musculoskeletal Science and Practice 2020 Feb;45:102103. [doi: 10.1016/j.msksp.2019.102103]
43. Shoemaker SJ, Wolf MS, Brach C. Development of the Patient Education Materials Assessment Tool (PEMAT): A new measure of understandability and actionability for print and audiovisual patient information. Patient Education and Counseling 2014 Sep;96(3):395-403. [doi: 10.1016/j.pec.2014.05.027]
44. Littlewood C, Malliaras P, Mawson S, May S, Walters SJ. Self-managed loaded exercise versus usual physiotherapy treatment for rotator cuff tendinopathy: a pilot randomised controlled trial. Physiotherapy 2014 Mar;100(1):54-60. [doi: 10.1016/j.physio.2013.06.001]
45. Mezirow J. Transformative Learning: Theory to Practice. New Directions for Adult and Continuing Education 2002 Dec 10;1997(74):5-12. [doi: 10.1002/ace.7401]
46. Perry J, VanDenKerkhof EG, Wilson R, Tripp DA. Development of a Guided Internet-based Psycho-education Intervention Using Cognitive Behavioral Therapy and Self-Management for Individuals with Chronic Pain. Pain Management Nursing 2017 Apr;18(2):90-101. [doi: 10.1016/j.pmn.2016.12.004]
47. Holmgren T, Bjornsson Hallgren H, Oberg B, Adolfsson L, Johansson K. Effect of specific exercise strategy on need for surgery in patients with subacromial impingement syndrome: randomised controlled study. BMJ 2012 Feb 20;344(feb20 1):e787-e787. [doi: 10.1136/bmj.e787]
48. Abraham C, Kelly MP, West R, Michie S. The UK National Institute for Health and Clinical Excellence public health guidance on behaviour change: a brief introduction. Psychol Health Med 2009 Jan;14(1):1-8. [doi: 10.1080/13548500802537903] [Medline: 19085307]
49. Bennell K, Wee E, Coburn S, Green S, Harris A, Staples M, et al. Efficacy of standardised manual therapy and home exercise programme for chronic rotator cuff disease: randomised placebo controlled trial. BMJ 2010 Jun 08;340:c2756 [FREE Full text] [doi: 10.1136/bmj.c2756] [Medline: 20530557]
50. Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Roijen LH. The iMTA Productivity Cost Questionnaire. Value in Health 2015 Sep;18(6):753-758. [doi: 10.1016/j.jval.2015.05.009]
51. Angst F, Schwyzer H, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function: Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH) and its short version (QuickDASH), Shoulder Pain and Disability Index (SPADI), American Shoulder and Elbow Surgeons (ASES) Society standardized shoulder assessment form, Constant (Murley) Score (CS), Simple Shoulder Test (SST), Oxford Shoulder Score (OSS), Shoulder Disability Questionnaire (SDQ), and Western Ontario Shoulder Instability Index (WOSI). Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S174-S188 [FREE Full text] [doi: 10.1002/acr.20630] [Medline: 22588743]
52. Roy J, MacDermid JC, Woodhouse LJ. Measuring shoulder function: A systematic review of four questionnaires. Arthritis Rheum 2009 May 15;61(5):623-632. [doi: 10.1002/art.24396]
53. Vicenzino B, Collins N, Crossley K, Beller E, Darnell R, McPoil T. Foot orthoses and physiotherapy in the treatment of patellofemoral pain syndrome: a randomised clinical trial. BMC Musculoskelet Disord 2008 Feb 27;9:27 [FREE Full text] [doi: 10.1186/1471-2474-9-27] [Medline: 18304317]

54. Page MJ, McKenzie JE, Green SE, Beaton DE, Jain NB, Lenza M, et al. Core domain and outcome measurement sets for shoulder pain trials are needed: systematic review of physical therapy trials. Journal of Clinical Epidemiology 2015 Nov;68(11):1270-1281. [doi: 10.1016/j.jclinepi.2015.06.006]
55. Lundberg MKE, Styf J, Carlsson SG. A psychometric evaluation of the Tampa Scale for Kinesiophobia — from a physiotherapeutic perspective. Physiotherapy Theory and Practice 2009 Jul 10;20(2):121-133. [doi: 10.1080/09593980490453002]
56. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: Development and validation. Psychological Assessment 1995;7(4):524-532. [doi: 10.1037/1040-3590.7.4.524]
57. Nicholas MK, McGuire BE, Asghari A. A 2-Item Short Form of the Pain Self-Efficacy Questionnaire: Development and Psychometric Evaluation of PSEQ-2. The Journal of Pain 2015 Feb;16(2):153-163. [doi: 10.1016/j.jpain.2014.11.002]
58. Hill J, Bird H. Patient knowledge and misconceptions of osteoarthritis assessed by a validated self-completed knowledge questionnaire (PKQ-OA). Rheumatology 2007 Mar 31;46(5):796-800. [doi: 10.1093/rheumatology/kel407]
59. Schünemann H. Interpreting results and drawing conclusions. In: Cochrane handbook for systematic reviews of interventions. Hoboken, NJ: Wiley-Blackwell; 2019:403-431.
60. Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. Pilot Feasibility Stud 2019 Oct 06;5(1). [doi: 10.1186/s40814-019-0499-1]
61. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are Health Behavior Change Interventions That Use Online Social Networks Effective? A Systematic Review. J Med Internet Res 2014 Feb 14;16(2):e40. [doi: 10.2196/jmir.2952]
62. Ryan J, Lopian L, Le B, Edney S, Van Kessel G, Plotnikoff R, et al. It's not raining men: a mixed-methods study investigating methods of improving male recruitment to health behaviour research. BMC Public Health 2019 Jun 24;19(1):814 [FREE Full text] [doi: 10.1186/s12889-019-7087-4] [Medline: 31234825]
63. Adamson J, Cockayne S, Puffer S, Torgerson DJ. Review of randomised trials using the post-randomised consent (Zelen's) design. Contemporary Clinical Trials 2006 Aug;27(4):305-319. [doi: 10.1016/j.cct.2005.11.003]
64. Geraghty AWA, Torres LD, Leykin Y, Perez-Stable EJ, Munoz RF. Understanding attrition from international internet health interventions: a step towards global eHealth. Health Promotion International 2012 Jul 10;28(3):442-452. [doi: 10.1093/heapro/das029]
65. Rosser B, Vowles KE, Keogh E, Eccleston C, Mountain GA. Technologically-assisted behaviour change: a systematic review of studies of novel technologies for the management of chronic illness. J Telemed Telecare 2009;15(7):327-338. [doi: 10.1258/jtt.2009.090116] [Medline: 19815901]
66. Bossen D, Veenhof C, Van Beek KE, Spreeuwenberg PM, Dekker J, De Bakker DH. Effectiveness of a Web-Based Physical Activity Intervention in Patients With Knee and/or Hip Osteoarthritis: Randomized Controlled Trial. J Med Internet Res 2013 Nov 22;15(11):e257. [doi: 10.2196/jmir.2662]
67. Wangberg SC, Bergmo TS, Johnsen JAK. Adherence in Internet-based interventions. Patient Prefer Adherence 2008 Feb 02;2:57-65 [FREE Full text] [Medline: 19920945]
68. Lambert TE, Harvey LA, Avdalis C, Chen LW, Jeyalingam S, Pratt CA, et al. An app with remote support achieves better adherence to home exercise programs than paper handouts in people with musculoskeletal conditions: a randomised trial. Journal of Physiotherapy 2017 Jul;63(3):161-167. [doi: 10.1016/j.jphys.2017.05.015]
69. Wickström G, Bendix T. The "Hawthorne effect"--what did the original Hawthorne studies actually show? Scand J Work Environ Health 2000 Aug;26(4):363-367 [FREE Full text] [Medline: 10994804]
70. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive System Design Does Matter: a Systematic Review of Adherence to Web-based Interventions. J Med Internet Res 2012 Nov 14;14(6):e152. [doi: 10.2196/jmir.2104]
71. Brooks MA, Beaulieu JE, Severson HH, Wille CM, Cooper D, Gau JM, et al. Web-based therapeutic exercise resource center as a treatment for knee osteoarthritis: a prospective cohort pilot study. BMC Musculoskelet Disord 2014 May 17;15(1):158 [FREE Full text] [doi: 10.1186/1471-2474-15-158] [Medline: 24884547]
72. Hopewell S, Keene DJ, Maia Schlüssel M, Dritsaki M, Dutton S, Carr A, et al. Clinical and cost-effectiveness of progressive exercise compared with best practice advice, with or without corticosteroid injection, for the treatment of rotator cuff disorders: protocol for a 2x2 factorial randomised controlled trial (the GRASP trial). BMJ Open 2017 Jul 17;7(7):e018004. [doi: 10.1136/bmjopen-2017-018004]
73. Ford I, Norrie J. Pragmatic Trials. N Engl J Med 2016 Aug 04;375(5):454-463. [doi: 10.1056/nejmra1510059]

Abbreviations

Aus $: Australian dollars
CERT: Consensus on Exercise Recording Template
CONSORT: Consolidated Standards of Reporting Trials
EQ5D: EuroQol 5D-5L
GP: general practitioner
GROC: Global Rating of Change

iPCQ: iMTA Productivity Cost Questionnaire
MCID: minimal clinically important difference
OA: osteoarthritis
PCS: Pain Catastrophizing Scale
PEMAT: Patient Education Materials Assessment Tool
PSEQ: Pain Self-Efficacy Questionnaire
RCRSP: rotator cuff–related shoulder pain
RCT: randomized controlled trial
SMD: standardized mean difference
SPADI: shoulder pain and disability index
TSK: Tampa Scale for Kinesiophobia
VAS: visual analog scale

Edited by L Buis, G Eysenbach; submitted 15.09.20; peer-reviewed by M Moessner, F Alburquerque-Sendin; comments to author 06.10.20; revised version received 14.10.20; accepted 21.10.20; published 18.11.20.

Please cite as:
Malliaras P, Cridland K, Hopmans R, Ashton S, Littlewood C, Page R, Harris I, Skouteris H, Haines T
Internet and Telerehabilitation-Delivered Management of Rotator Cuff–Related Shoulder Pain (INTEL Trial): Randomized Controlled Pilot and Feasibility Trial
JMIR Mhealth Uhealth 2020;8(11):e24311
URL: http://mhealth.jmir.org/2020/11/e24311/
doi:10.2196/24311
PMID:33206059

Review

Implementation of Telerehabilitation Interventions for the Self-Management of Cardiovascular Disease: Systematic Review

Narayan Subedi[1], MPH; Jonathan C Rawstorn[1], PhD; Lan Gao[2], PhD; Harriet Koorts[1], PhD; Ralph Maddison[1], PhD

[1]School of Exercise and Nutrition Sciences, Faculty of Health, Deakin University, Melbourne, Australia

[2]School of Health and Social Development, Faculty of Health, Deakin University, Melbourne, Australia

Corresponding Author:
Narayan Subedi, MPH
School of Exercise and Nutrition Sciences
Faculty of Health
Deakin University
221 Burwood Highway
Burwood Victoria
Melbourne, 3125
Australia
Phone: 61 404745397
Email: nsubedi@deakin.edu.au

Abstract

Background: Coronary heart disease (CHD) is a leading cause of disability and deaths worldwide. Secondary prevention, including cardiac rehabilitation (CR), is crucial to improve risk factors and to reduce disease burden and disability. Accessibility barriers contribute to underutilization of traditional center-based CR programs; therefore, alternative delivery models, including cardiac telerehabilitation (ie, delivery via mobile, smartphone, and/or web-based apps), have been tested. Experimental studies have shown cardiac telerehabilitation to be effective and cost-effective, but there is inadequate evidence about how to translate this research into routine clinical practice.

Objective: This systematic review aimed to synthesize research evaluating the effectiveness of implementing cardiac telerehabilitation interventions at scale in routine clinical practice, including factors underlying successful implementation processes, and experimental research evaluating implementation-related outcomes.

Methods: MEDLINE, Embase, PsycINFO, and Global Health databases were searched from 1990 through November 9, 2018, for studies evaluating the implementation of telerehabilitation for the self-management of CHD. Reference lists of included studies and relevant systematic reviews were hand searched to identify additional studies. Implementation outcomes of interest included acceptability, appropriateness, adoption, feasibility, fidelity, implementation cost, penetration, and sustainability. A narrative synthesis of results was carried out.

Results: No included studies evaluated the implementation of cardiac telerehabilitation in routine clinical practice. A total of 10 studies of 2250 participants evaluated implementation outcomes, including acceptability (8/10, 80%), appropriateness (9/10, 90%), adoption (6/10, 60%), feasibility (6/10, 60%), fidelity (7/10, 70%), and implementation cost (4/10, 40%), predominantly from the participant perspective. Cardiac telerehabilitation interventions had high acceptance among the majority of participants, but technical challenges such as reliable broadband internet connectivity can impact acceptability and feasibility. Many participants considered telerehabilitation to be an appropriate alternative CR delivery model, as it was convenient, flexible, and easy to access. Participants valued interactive intervention components, such as real-time exercise monitoring and feedback as well as individualized support. The penetration and sustainability of cardiac telerehabilitation, as well as the perspectives of CR practitioners and health care organizations, have received little attention in existing cardiac telerehabilitation research.

Conclusions: Experimental trials suggest that participants perceive cardiac telerehabilitation to be an acceptable and appropriate approach to improve the reach and utilization of CR, but pragmatic implementation studies are needed to understand how interventions can be sustainably translated from research into clinical practice. Addressing this gap could help realize the potential impact of telerehabilitation on CR accessibility and participation as well as person-centered, health, and economic outcomes.

Trial Registration: International Prospective Register of Systematic Reviews (PROSPERO) CRD42019124254; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=124254

(JMIR Mhealth Uhealth 2020;8(11):e17957) doi:10.2196/17957

KEYWORDS

heart diseases; cardiac rehabilitation; telerehabilitation; implementation science; smartphone; systematic review

Introduction

Cardiovascular diseases (CVDs) are a leading cause of clinical (ie, death and disability), health, and economic burden globally, accounting for approximately 31% (17.9 million) of total deaths each year [1-3]. Coronary heart disease (CHD), including myocardial infarction (MI) and angina, is the most common and burdensome form of CVD [4,5]. CHD accounts for a high proportion of all CVD deaths and more disability-adjusted life years than diseases such as cancer and diabetes [4-6]. Therefore, secondary prevention interventions that support CVD management are critical to reducing disease burden and health care expenditure.

Cardiac rehabilitation (CR) is an essential component of secondary prevention for CHD that comprises coordinated, multifaceted interventions designed to improve physical, psychological, and social functioning [7-12]. CR includes medical evaluation, exercise prescription, cardiac risk factor modification, education, and counseling [13]. CR is safe, effective [14], and more cost-effective than no CR on overall health service expenditure [15-17]. Systematic reviews have shown that participation in center-based programs (ie, face-to-face delivery) reduces risks of hospital admissions and cardiac mortality, and improves health-related quality of life [14,18].

Despite these benefits, uptake and adherence of center-based CR are suboptimal [19-21]. Reasons for this are multifaceted [22-26], but accessibility-related factors, such as limited availability of programs, transportation, and parking, are prominent [22,24-28]. For these reasons, home-based delivery models have been tested to improve access and participation outside of clinical settings [29].

Home-based CR, which typically includes print resources, home visits, and/or telephone calls, has been shown to be as effective as center-based programs for improving health-related quality of life, CVD risk factors, and mortality [30]. However, few CR services offer home-based options (eg, less than one-quarter in the United Kingdom, United States, and Australia [30,31]). In addition, home-based programs are typically unable to provide the level of supervision, individualized coaching, and feedback from CR professionals that is common in center-based programs. Therefore, alternative delivery models that combine the accessibility of home-based programs with the comprehensive support of center-based CR are needed.

The use of information and communication technologies (ICTs) to connect participants and CR professionals, which is termed cardiac telerehabilitation [32], has been investigated as an alternative. Systematic reviews have demonstrated the effectiveness of cardiac telerehabilitation for improving cardiovascular risk factors and health-related quality of life [33-37]. However, early telerehabilitation interventions were mostly limited to telephone counseling, which limits the types of rehabilitation support that can be provided [38]. Technological innovations including mobile phones, particularly smartphones, and mobile broadband [39,40] have enabled more flexible cardiac telerehabilitation interventions [36,41-43].

Recent studies using cutting-edge technologies, such as smartphones, mobile apps, and the internet, have demonstrated that cardiac telerehabilitation can deliver more comprehensive services [44], including individualized real-time exercise monitoring and coaching, similar to center-based programs [45]. Growing evidence indicates telerehabilitation could substantially broaden the benefits and impact of CR; however, most interventions have only been evaluated in controlled experimental settings (eg, [41-45]). There is little evidence to guide the successful, scalable, sustainable translation of telerehabilitation into real-world settings [46-48]; that is, there is a lack of studies that have tested telerehabilitation interventions when delivered by health care staff in routine clinical practice.

Real-world implementation of an intervention is contextually dependent, influenced by individual (ie, personal characteristics), organizational (ie, hospital or service organization), community (ie, local government), and system-level (ie, government) factors, all of which are difficult to control in experimental designs [49]. Many public health interventions fail to be adopted or are less likely to be scaled and sustained when delivered in real-world settings, and the complexities and challenges involved in real-world implementation and scale-up are partly responsible for this lack of translational success [50]. A greater understanding of factors related to the implementation of interventions in practice settings is imperative for increasing population-level impact [51].

The purpose of this review was to synthesize research evaluating the implementation of cardiac telerehabilitation interventions when delivered in routine clinical practice.

Methods

Registration

This review was registered in PROSPERO (International Prospective Register of Systematic Reviews) (CRD42019124254) before screening search results, and was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [52,53].

Information Sources and Search Strategy

Electronic databases—MEDLINE, Embase, PsycINFO, and Global Health—were searched between January 1990 and November 9, 2018, for studies that combined three concepts: telehealth, CVD, and implementation science. The search strategy was created for MEDLINE and modified for the other databases (see Multimedia Appendix 1).

Eligibility and Study Selection

Eligible studies were those that evaluated the implementation of cardiac telerehabilitation in routine clinical practice or assessed implementation outcomes of interest in experimental studies, including randomized and nonrandomized designs, among adults (aged ≥18 years) with CHD (ie, MI, angina, and coronary revascularization).

Cardiac telerehabilitation interventions were defined as those with at least 50% of the program delivered via ICT, including any mobile phone (ie, feature phone or smartphone), web-based platforms, or wireless devices such as sensors. Implementation in routine clinical practice was defined as interventions delivered as part of existing CR services, without significant ongoing input from a research team.

Experimental studies were included as we anticipated few studies examining real-world implementation, and experimental studies provide the next best available evidence to advance the field; eligibility was not limited to randomized controlled trials (RCTs) to allow the inclusion of translational studies that used alternative study designs. In addition to other criteria, eligible experimental studies were those that assessed constructs defined in a taxonomy of key constructs related to the effective implementation of evidence-based interventions, including acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration, and sustainability [54]. To acknowledge the importance of multiple stakeholder levels in successful implementation projects [55,56] and to meet the aims of this review, we assessed these constructs at the consumer (ie, participant), individual provider (ie, CR practitioner), and provider (ie, health care organization or institution) levels [54]. Constructs were defined as follows:

1. Acceptability: satisfaction among implementation stakeholders with different aspects of the intervention, such as content, delivery, and complexity. Stakeholders included health care consumers, practitioners, and health care organization operational staff who participated in, delivered, and oversaw the provision of CR services, respectively.
2. Adoption: rates of uptake or utilization of the intervention at the practitioner and/or health care organization level.
3. Appropriateness: program suitability or compatibility at the health care consumer, practitioner, and/or health care organization level.
4. Feasibility: practicability of the intervention for everyday use at the practitioner and/or health care organization level.
5. Fidelity: delivery of the intervention as designed.
6. Implementation cost: assessments of marginal cost, cost-effectiveness, or cost benefit.
7. Penetration: the degree to which the intervention was institutionalized within health care organizations.
8. Sustainability: continued delivery of the intervention beyond the study period, as well as characteristics of the implementation context that did or could influence the continuation of intervention delivery [54].

Feasibility and pilot studies were excluded from this review. To meet the aims of this review, it was important to include only interventions that had already undergone preliminary testing for feasibility and were considered by their respective authors as feasible for testing in the trial or delivery in practice. Conference abstracts, nonhuman studies, non-English-language papers, and grey literature were also excluded. Systematic reviews and study protocols were not eligible for inclusion; however, relevant systematic reviews were searched for eligible studies and cited where appropriate, and results articles were sought for relevant study protocols.

Search results were exported to a reference manager, EndNote X8 (Clarivate), for duplicate removal, then transferred to Rayyan (Qatar Computing Research Institute) for screening [57]. Records were assessed by NS, verified by JR and HK, and underwent full-text review if the title or abstract identified the specified population and intervention components.

Data Extraction

Data describing eligibility, study design, participant and intervention characteristics, risk of bias, and outcomes of interest were extracted by NS using a standardized electronic form and verified by JR.

Risk-of-Bias Assessment

Risks of selection, performance, detection, attrition, reporting, and other biases in included experimental studies were assessed using the Cochrane risk-of-bias tool [58]. Risks of bias in nonrandomized cohort studies were assessed using the Joanna Briggs Institute Critical Appraisal Checklist for Cohort Studies [59]. Risks of bias were assessed by NS and verified by JR. When available, risk-of-bias assessments were augmented with study protocols and clinical trial registrations.

Data Synthesis

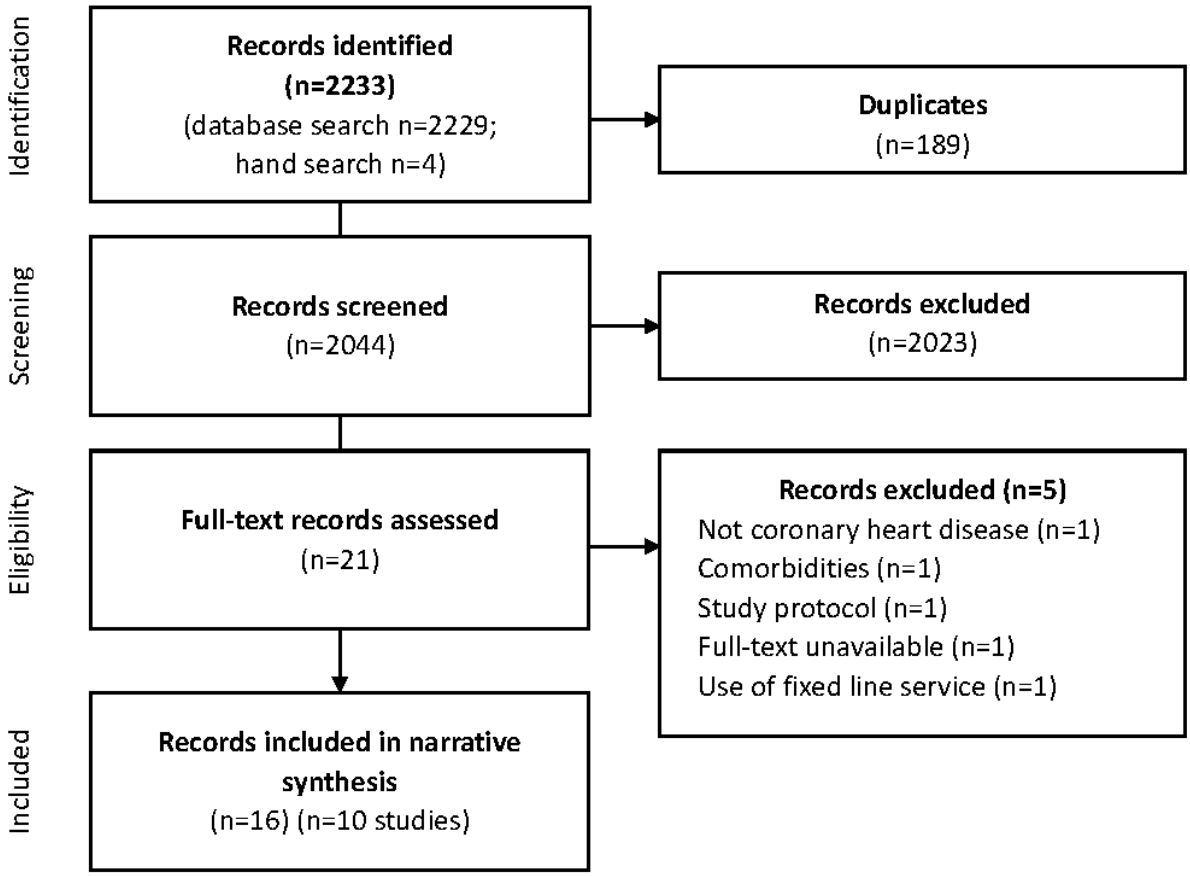
A narrative synthesis of the data was carried out in this review.

Results

Study Selection

In total, 2044 unique study reports were screened. From these, 21 underwent full-text review; 16 reports describing 10 studies (2250 participants in total) met the eligibility criteria and were included in the narrative synthesis [41-45,60-64]. Study selection is summarized in Figure 1.

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



Characteristics of Included Studies

Included studies were published between 2013 and 2018; studies were conducted in developed countries, including New Zealand [42,43,45], Australia [41,44], Canada [61], the Netherlands [60], Poland [63], and the United States [62]. One multi-country study was carried out in Spain, Germany, and the United Kingdom [64].

No eligible studies were identified that evaluated the effectiveness of the implementation of cardiac telerehabilitation in routine clinical practice. Of the 10 included studies, 9 (90%) used randomized controlled experimental designs [41-45,60-62,64], while 1 (10%) used an uncontrolled pre-post intervention design [63].

The mean age of study participants ranged from 55 to 65 years, and most participants were male (72%-93%). All studies recruited participants via hospitals or community-based CR centers. Detailed characteristics of the included studies (ie, study design, treatments, and primary and implementation outcomes) are presented in Multimedia Appendix 2.

Cardiac Telerehabilitation Intervention Characteristics

Out of 10 studies, 3 interventions (30%) were delivered using a mobile phone, smartphone, or web-based platform alone [41,61,62], while remaining interventions used combinations of web-based content, mobile phones or smartphones, and sensors. The most commonly targeted lifestyle risk factors were physical activity, diet, tobacco smoking, and medication adherence. Intervention duration ranged from 30 days [62] to 24 weeks [41-43] (see Multimedia Appendix 2).

Most interventions (6/10, 60%) comprised a messaging component (eg, SMS or push notifications) [41,43-45,61,62] to educate or motivate participants to improve self-management behaviors. Out of 10 interventions, 7 (70%) enabled communication between providers and participants via a web-based program, mobile phone or smartphone, and/or telephone [42,44,45,60,61,63,64]. Out of 10 interventions, 6 (60%) included exercise monitoring [42,44,45,60,63,64], including 2 (20%) that provided live guidance [64] or real-time monitoring and coaching during exercise [45]. Out of 10 studies, 5 (50%) delivered telerehabilitation in combination with usual care (ie, center- or community-based CR) [41-43,45,60], 4 (40%) delivered cardiac telerehabilitation alone [44,61,62,64], and 1 (10%) delivered a hybrid intervention comprising center-based and telerehabilitation components [63] (see Multimedia Appendix 2).

Risk of Bias in the Included Studies

The quality of the included studies in the review varied (see Multimedia Appendix 2). Risk of bias was judged to be low in 6 out of 10 (60%) experimental studies [41-43,45,60,61] and high in 3 (30%) studies [44,62,64]. High risk of bias was judged due to incomplete outcome data and lack of blinding of participants and outcomes.

The single nonrandomized study (1/10, 10%) had a high risk of bias due to lack of a control group, not identifying

confounding factors, and inadequate reporting of follow-up time [63].

Implementation Outcomes

Overview

Included studies reported between three and six implementation outcomes. Appropriateness, acceptability, fidelity, adoption and feasibility, were assessed in 9 (90%), 8 (80%), 7 (70%), and 6 (60%), of the 10 studies, respectively; cost of intervention was assessed in only 4 (40%) studies, and penetration and sustainability were not assessed. Outcomes were predominantly assessed from a participant perspective, rather than from individual provider (ie, practitioner) or organizational perspectives. Implementation outcome findings are summarized below, with supporting data provided in Table 1 [41-45,60-71].

Table 1. Implementation outcomes for telerehabilitation interventions.

Study author, year, and implementation construct	Implementation outcomes
Chow, 2015 [41,65]	
Acceptability	SMS intervention acceptability was 90.9% (279/307); request to stop SMS was 2.3% (7/307)
Adoption	Focus groups reported high user engagement with saving and sharing SMS messages, receiving support from providers and family, and message personalization
Appropriateness	SMS was useful: 90.9% (279/307) SMS was easy to understand: 96.7% (297/307) SMS was motivating for change: 77.2% (237/307); especially for diet (249/307, 81.1%), exercise (223/307, 72.6%), and medication adherence (234/307,76.2%) Appropriateness of language used in SMS: 94.8% (291/307) Appropriateness of SMS frequency (4 times/week): 86.0% (264/307); timing : 89.9% (276/307, random timing was considered ideal); and 6-month duration: 77.2% (237/307)
Feasibility	Not assessed
Fidelity	96.0% (338/352) of participants received all scheduled messages (analytic data) and read ≥75% of SMS messages: 95.4% (293/307 self-report survey respondents)
Implementation cost	US $0.10/SMS message (<US $10 per capita)
Penetration	Not assessed
Sustainability	Not assessed
Dale, 2015 [42]	
Acceptability	Satisfaction with 24-week program duration was 79% (48/61) and with number of SMS messages was 84% (51/61) Recommend to other people: 90% (55/61)
Adoption	98% (60/61) of participants initiated the SMS intervention ≥1 website login: 75% (46/61); median 3, range 0-100
Appropriateness	90% (55/61) and 43% (26/61) of participants felt that SMS messages and the website were good cardiac rehabilitation (CR) delivery methods, respectively Appropriate number of SMS messages: 84% (51/61) Intervention useful for learning about (47/61, 77%) and recovering from (51/61, 84%) a heart event and for changing behaviors, such as physical activity (39/61, 64%) and consumption of fruit and vegetables (37/61, 61%), saturated fat (34/61, 56%), and salt (26/61, 43%)
Feasibility	Not assessed
Fidelity	Read all SMS messages: 85% (52/61) Sent ≥1 SMS step count message: 95% (58/61); mean of 15 submissions (SD 8.7) over 24 weeks
Implementation cost	Not assessed
Penetration	Not assessed
Sustainability	Not assessed
Kraal, 2013 [60,66]	
Acceptability	Satisfaction was higher for telerehabilitation than for center-based rehabilitation (8.7/10 vs 8.1/10; *P*=.02)
Adoption	Not assessed
Appropriateness	Not assessed
Feasibility	Not assessed
Fidelity	Exercise adherence was similar in telerehabilitation and center-based rehabilitation (mean 22.0, SD 6.8, vs mean 20.6, SD 4.3, sessions)
Implementation cost	Similar per-capita cost to deliver telerehabilitation and center-based rehabilitation (€314 vs €336) Per-capita costs did not differ between telerehabilitation and center-based rehabilitation for total health care use (mean €2419, SD 1968, vs mean €2855, SD 2797; *P*=.39) or total work absenteeism (mean €3846, SD 8400, vs mean €6569, SD 8170; *P*=.12) Probability of cost-effectiveness was higher for telerehabilitation than for center-based rehabilitation under several assumptions

Study author, year, and implementation construct	Implementation outcomes
Penetration	Not assessed
Sustainability	Not assessed
Lear, 2015 [61,67]	
Acceptability	22 purposively sampled interviews reported satisfaction, acceptability, and confidence in using virtual CR
Adoption	High self-reported engagement and utilization in virtual CR (interview data) Mean website log-ins was 27 per participant (range 0-140) Mean engagement in chat sessions with health care providers was 3.6
Appropriateness	Virtual CR perceived to be accessible and effective
Feasibility	Virtual CR perceived to be convenient
Fidelity	Not assessed
Implementation cost	Not assessed
Penetration	Not assessed
Sustainability	Not assessed
Maddison, 2015 [43,68]	
Acceptability	SMS and website intervention components were liked by 57% (43/75) and 73% (55/75) of participants, respectively Acceptability of 24-week intervention duration: 71% (53/75) Acceptability of message delivery timing: 57% (43/75); exercise ideas SMS content: 77% (58/75); exercise benefits education content: 79% (59/75); and website content: 47% (35/75); 49% (37/75) did not use the website
Adoption	Not assessed
Appropriateness	Some (number not reported) participants who were already exercising felt the intervention was unnecessary or the exercise prescription was not relevant
Feasibility	Difficulties using website: 17% (13/75) Major barriers were lack of high-speed broadband or knowledge about using websites
Fidelity	93% (70/75) read most SMS messages 64% (48/75) used the website (visits per participant: mean 11, SD 16, range 0-82)
Implementation cost	NZ $239 per capita (intervention set-up + delivery only; health care utilization and indirect societal costs excluded) Incremental cost-effectiveness ratio: NZ $28,768 per quality-adjusted life year (QALY) Probability of cost-effectiveness: 72% (willingness to pay: NZ $20,000 per QALY) and 90% (willingness to pay: NZ $50,000 per QALY)
Penetration	Not assessed
Sustainability	Not assessed
Maddison, 2019 [45,69]	
Acceptability	87% (58/67) would choose telerehabilitation instead of center-based rehabilitation if implemented in clinical practice Satisfaction with individualized exercise prescription: 90% (60/67); real-time exercise monitoring: 94% (63/67); encouragement and social support: 87% (58/67); behavior change messages: 85% (57/67); self-monitoring: 96% (64/67); and goal-setting features: 69% (46/67)
Adoption	94% (77/82) of participants initiated telerehabilitation
Appropriateness	97% (65/67) of patients reported that telerehabilitation is a good approach for delivering exercise rehabilitation
Feasibility	Wearable sensor is easy to use: 99% (66/67); and is comfortable: 97% (65/67) Smartphone app is easy to use: 79% (53/67); easy to understand: 87% (58/67); and reliable: 66% (44/67) Rare technical difficulties, commonly solved with familiarization

Study author, year, and implementation construct	Implementation outcomes
Fidelity	Adherence to prescribed exercise was comparable in telerehabilitation (mean 58.34%, SD 36.58, range 0-100) and center-based rehabilitation (mean 63.80%, SD 30.59, range 0-100; *P*=.31)
Implementation cost	Lower per-capita program delivery cost for telerehabilitation than for center-based rehabilitation (NZ $1130 vs NZ $3466)
	No difference in total (ie, program delivery + health care and medication utilization) per-capita cost (NZ $4920 vs NZ $9535)
Penetration	Not assessed
Sustainability	Not assessed
Park, 2014 [62]	
Acceptability	Strong or moderate agreement about intervention satisfaction: 82% (23/28) for SMS reminders + education; and 88% (22/25) for SMS education alone
Adoption	Not assessed
Appropriateness	Strong or moderate agreement that the interventions were useful for assisting medication adherence: 71% (20/28) for SMS reminders + education; and 48% (12/25) for SMS education alone
Feasibility	Strong or moderate agreement that interventions were easy to use: 88.6%
	Technical difficulties receiving SMS: 7.6%
Fidelity	Not assessed
Implementation cost	Not assessed
Penetration	Not assessed
Sustainability	Not assessed
Piotrowicz, 2014 [63,70]	
Acceptability	Not assessed
Adoption	Not assessed
Appropriateness	Felt safer during exercise with hybrid telerehabilitation than unsupervised: 80.9%
	Hybrid telerehabilitation was useful for increasing exercise: 95%; daily physical activity: 80%; and mental health: 71%
Feasibility	Telemonitoring device was very easy or easy to use: 98.3%
	No problems self-fitting electrocardiogram (ECG) electrodes: 99.4%
	No problems transmitting ECG from home: 84%
	Missed ≥1 exercise session due to technical difficulties: 39.3%
	Problems communicating with telemonitoring center: 62.8%
Fidelity	Not assessed
Implementation cost	Not assessed
Penetration	Not assessed
Sustainability	Not assessed
Salvi, 2018 [64,71]	
Acceptability	Guided exercise telerehabilitation ratings (mean [95% CI] rating score, max 5) for ease of use: 3.53 (2.94-4.12); interest: 4.42 (4.11-4.74); stimulation: 3.95 (3.49-4.41); and enjoyment: 3.84 (3.46-4.22)
	nb: data represent only 35% (19/55) of participants randomized to telerehabilitation
Adoption	73% (40/55) of participants initiated guided exercise telerehabilitation
	Nonadoption was attributed to unavailability of the clinical team
Appropriateness	Guided exercise telerehabilitation ratings (mean [95% CI] rating score, max 5) for usefulness to increase motivation: 4.59 (4.35-4.83); to increase safety: 4.47 (4.13-4.81); and to increase compliance: 4.47 (3.93-5.01)
	Overall, guided exercise telerehabilitation was considered appropriate for its purpose

Study author, year, and implementation construct	Implementation outcomes
Feasibility	Exercise sessions affected by technical errors: 18% (ie, poor biosensor signal or connectivity and poor transmission of data to server)
	Suboptimal internet connectivity prevented 15 participants from recording or completing any exercise sessions
	6 dropouts were attributed to technical challenges
Fidelity	Participants initiated (mean [95% CI]) 61% (76%-46%) of the prescribed number of exercise sessions (79% [91%-67%] among 17 participants who completed the study) and completed 32% (44%-20%) of the prescribed duration of exercise (45% [59%-31%] among 17 participants who completed the study)
Implementation cost	Not assessed
Penetration	Not assessed
Sustainability	Not assessed
Varnfield, 2014 [44]	
Acceptability	Not assessed
Adoption	Program uptake (ie, completion of ≥1 exercise session) was higher in telerehabilitation than center-based rehabilitation: 80% (48/60) vs 62% (37/60); relative risk (RR)=1.30, 95% CI 1.03-1.64; *P*<.05
Appropriateness	Smartphone-measured step counts increased motivation to reach exercise goals: 84% (38/45)
Feasibility	Not assessed
Fidelity	Categorical adherence (ie, completing 4/6 weeks of exercise training) was higher in telerehabilitation than center-based rehabilitation: 95% (45/48) vs 68% (25/37); RR=1.40, 95% CI 1.13-1.70; *P*<.05
Implementation cost	Not assessed
Penetration	Not assessed
Sustainability	Not assessed

Acceptability

Out of 10 included studies, 8 (80%) [41-43,45,60-62,64] reported the acceptability of telerehabilitation interventions from the participant perspective only; none reported acceptability from the individual provider or organization perspectives. Specific outcome measures within the acceptability implementation construct included perceived acceptability, satisfaction, likes and dislikes, interest, stimulation, and enjoyment. Studies reported high rates of acceptance for cardiac telerehabilitation, ranging from 71% [43] to 99% of participants [45]. Interventions that facilitated interaction between participants and providers [42,45,60,61,64] and delivered individually tailored content [43,60], in particular, appeared to have high acceptability. However, 4 studies out of 10 (40%) reported lack of interest among some participants [43,62,64]. In particular, messaging interventions (eg, SMS and push notifications) were not satisfactory for participants who would prefer face-to-face interaction with rehabilitation professionals [43]. Usability challenges such as insufficient internet connectivity, which fall within the *feasibility* implementation construct (see Feasibility section below), can also impact negatively on acceptability [42,64].

Adoption

Out of 10 included studies, 6 (60%) reported adoption of the intervention from the participant perspective only; none reported adoption from the individual provider or organization perspectives [41,42,44,45,61,64]. Specific outcome measures within the adoption implementation construct included initial uptake and engagement with telerehabilitation interventions. While levels of adoption varied between studies, they were generally high across SMS, website, and smartphone-based interventions. Out of 10 studies, 1 (10%) attributed some lack of adoption to the availability of staff delivering the intervention [64].

Appropriateness

Out of 10 included studies, 9 (90%) reported on the appropriateness of cardiac telerehabilitation from the participant perspective; none reported on appropriateness from the individual provider or organization perspectives [41-45,61-64]. Studies included a very broad range of outcome measures within the appropriateness implementation construct, including usefulness; suitability as an alternative CR delivery model; perceptions of safety, reassurance, accessibility, and effectiveness; as well as appropriateness of the intervention content, language, frequency, and duration.

These outcomes were positively appraised by the majority of participants, and cardiac telerehabilitation was perceived as convenient, flexible, safe, instant, private, and user-friendly. Many participants considered messaging (eg, SMS and push notifications) and smartphone apps to be appropriate mechanisms for delivering CR support [41,42,45,61,63]. Comprehensibility and interactivity also appeared to support participants' perceptions of intervention appropriateness. [41-43,45,60,61,63,64]. Participants reported that telerehabilitation interventions were useful for increasing motivation and confidence to exercise, modifying health

behaviors such as medication adherence and healthy eating, self-monitoring their health condition, and facilitating remote access to individualized exercise support [41,42,45,61,62,64]. Small proportions of participants perceived SMS to be either inadequate or unnecessary [41-43,62]. Overall, cardiac telerehabilitation interventions were considered appropriate by most participants [41-45,61-64]; however, significant heterogeneity of intervention designs and outcome measures in the literature we reviewed makes it difficult to identify factors that optimize intervention appropriateness.

Feasibility

Out of 10 included studies, 6 (60%) reported aspects of feasibility from a participant perspective but not from the individual provider or organization perspectives. Specific outcome measures within the feasibility implementation construct included usability, suitability of interventions for everyday use among participants, system reliability, or technical difficulties experienced by participants [43,45,61-64]. No included studies assessed the feasibility of telerehabilitation delivery from individual provider or organization perspectives.

Large majorities of telerehabilitation participants self-reported that technologies such as SMS, wearable sensors, smartphone apps, and websites were easy to use, convenient, comfortable, and easy to understand [43,45,61-63]. However, some technical challenges were noted. Approximately 20%-30% of participants reported reliability issues during real-time, remotely monitored, exercise rehabilitation, although the authors did not report whether issues were related to the required smartphone app, wearable sensor, or broadband internet connection [45]. Out of 10 studies, 4 (40%) reported a negative impact of unreliable broadband connectivity on user experiences during interventions that included web-based components and/or transmission of data to CR providers, which, at worst, can prevent participants from initiating their telerehabilitation intervention at all [43,62-64].

Fidelity

Out of 10 included studies, 7 (70%) reported on the fidelity of intervention receipt and/or completion among participants [41-45,60,64]. Specific outcome measures within the fidelity implementation construct included participant responsiveness or adherence to the intervention, such as receiving all the scheduled SMS messages, program completion, and adherence to prescribed exercise or medication. Large majorities (≥75%) of participants in SMS interventions self-reported reading all or most messages [41-43]. However, self-reported use of an intervention website was lower [43]. Out of 10 studies, 3 (30%) demonstrated that adherence to prescribed exercise telerehabilitation sessions was comparable to [45,60], or better than [44], center-based comparators.

Telerehabilitation interventions appeared to be delivered as per study protocols, with the exception of deviations caused by technical challenges (see Feasibility section above); however, only 1 study out of 10 (10%) formally assessed the fidelity of intervention delivery; analytic data indicated all SMS messages were successfully delivered to 96% of participants [41].

Implementation Cost

Only 4 of 10 (40%) included studies reported analyses of telerehabilitation intervention cost [41,43,45,60]. As interventions appear to have been provided at no cost to participants, these data likely represent cost from an organizational perspective. Specific outcome measures within this implementation construct included intervention delivery cost and cost-effectiveness. Intervention delivery costs varied markedly from US $10 per capita for an SMS intervention [41] to NZ $1130 per capita for a smartphone-based intervention that delivered real-time remote exercise supervision and coaching [45]. Out of 10 studies, 1 (10%) reported that the telerehabilitation intervention delivery cost was comparable to center-based CR [60], while another (1/10, 10%) reported almost 70% lower delivery costs for telerehabilitation compared with center-based programs [45]. Out of 10 studies, 2 (20%) that conducted cost-effectiveness analyses reported a 72%-90% probability of cost-effectiveness for an SMS intervention, assuming willingness-to-pay thresholds of NZ $20,000-$50,000 per quality-adjusted life year [43], and moderate to high probabilities that telerehabilitation would be more cost-effective than center-based rehabilitation, particularly at low willingness-to-pay thresholds [60]. None of the studies included in this review evaluated how telerehabilitation could be funded as an adjunct to existing CR services (ie, in additional to center-based program delivery costs).

Penetration and Sustainability

No included studies assessed any outcome measures within the penetration or sustainability implementation constructs.

Discussion

Principal Findings

The primary finding of our review is that, despite encouraging evidence for effectiveness [33-37], there is a lack of evidence evaluating the translation of cardiac telerehabilitation interventions from research into routine clinical practice. It is unclear whether this suggests that evidence-based interventions have yet to be implemented in clinical practice, have been implemented without evaluation, or have been implemented and evaluated but not yet published in the scientific literature.

The next best available data comes from a small number of experimental studies that have assessed key constructs related to the effective implementation of evidence-based interventions. Almost all included studies reported factors related to intervention appropriateness, acceptability, and fidelity; however, adoption, fidelity, and cost have received less attention, and intervention penetration and sustainability had yet to be evaluated.

Moreover, while consumers (ie, CR participants), individual providers (ie, CR practitioners), and organizations (ie, health care services) all play critical roles in achieving successful implementation outcomes [54,72], the cardiac telerehabilitation literature we reviewed has focused only on the consumer perspective. This may reflect the lack of research conducted in routine clinical practice, as individual and organizational providers may have little involvement in the delivery or

management of telerehabilitation interventions during experimental trials.

Cardiac telerehabilitation was generally well accepted among the majority of participants, even across a broad range of different interventions. While the small number of studies in our review makes it difficult to determine which interventions may be most acceptable to participants, intervention features that enable participants to communicate with practitioners and receive tailored or individualized support appear to promote high rates of acceptance. Unfortunately, acceptability has not yet been evaluated from a delivery perspective, so the perceptions of rehabilitation providers (ie, individual practitioners and organizations) remain unknown. Cardiac telerehabilitation was considered an appropriate delivery model by many participants, particularly those who value convenient, flexible, and accessible intervention support. Moreover, many participants reported that cardiac telerehabilitation was useful for improving their self-management of lifestyle behaviors and CVD risk factors. At the participant level, acceptability and appropriateness also appear to be moderated by intervention feasibility. Interventions that were simple to access, easy to use, reliable, and delivered through ubiquitous mobile, smartphone, and/or web technologies appeared to have higher acceptability and appropriateness.

As our review findings are drawn from experimental studies, it is unclear if they would generalize to the delivery of cardiac telerehabilitation in routine clinical practice. In particular, the predominance of randomized treatment allocation in included studies differs markedly from a recommendation that CR participants should be offered a choice of alternative CR delivery models that best fit their needs and preferences [73]. Studies that include preference-based treatment allocation are needed to mimic this key element of routine clinical practice or, better yet, evaluation studies should be conducted in parallel with the translation of cardiac telerehabilitation into routine clinical practice. The single non-RCT included in the review reported high acceptance and usability among participants who preferred telerehabilitation [63].

Reporting of participant-level feasibility was mixed, which may reflect the difficulty of documenting both the incidence and impact of usability and technical challenges. However, collectively the evidence suggests telerehabilitation is feasible for most participants. Feasibility among individual and organizational providers is also critical for successful implementation [74,75] but, similar to other implementation constructs, was not evaluated in the included studies.

Fidelity is one of the important implementation outcomes, as it contributes to intervention quality [54]. While 7 of 10 (70%) included studies reported high intervention uptake, some outcome measures of adherence may be confounded by self-reporting bias. Reassuringly, 3 (30%) studies indicate that adherence to cardiac telerehabilitation can be at least as high as center-based programs [44,45,60]. A key gap in the literature we reviewed is the lack of information about the fidelity of intervention delivery from the perspectives of individual and organizational providers, which is critical to understand when implementing new interventions to maintain intervention quality [76]. This may reflect known challenges in measuring implementation constructs and a lack of available validated tools [54]; however, we note that the delivery fidelity of interventions that require little, if any, provider input (eg, SMS) may be sufficiently assessed via software analytics.

There was little specific evidence that intervention delivery deviated from study protocols in the included studies, but it is unclear whether this indicates high fidelity intervention delivery or a lack of documentation to support such a conclusion. While future translational research could comprehensively evaluate the fidelity of intervention delivery, comparisons with preceding experimental research may be confounded by assumptions about equivalence between experimental and translational research contexts.

Intervention cost is a key contributor to low uptake of interventions by health care providers [77], and comparison of costs and cost-effectiveness is crucial for making evidence-based decisions about implementing and scaling new interventions [76]. Relatively few included studies reported economic analyses, and telerehabilitation costs varied markedly across different intervention designs. Our review indicates that telerehabilitation can reduce CR delivery costs and be cost-effective, but it is unclear if cost-effectiveness varies between different types of interventions. For example, interventions that require significant practitioner input, such as real-time remote exercise monitoring and coaching [45], may be substantially cheaper to deliver than center-based programs but more expensive than semiautomated interventions, such as SMS [43]. Whether or not an intervention represents good value depends on health effects and costs, as well as intervention and health care objectives. Therefore, examining the cost and health effects of telerehabilitation interventions in routine clinical practice is essential to provide more valuable information for implementation and scale-up.

The remaining implementation constructs of interest in this review—penetration and sustainability—were not assessed in any included studies. This was not surprising, as penetration and sustainability are more relevant to the mid- and later stages of implementation [54]. Short-duration experimental trials limit penetration to those who are willing to volunteer for research and preclude assessment of longer-term intervention sustainability. As a result, we lack valuable information on factors that could influence the sustainability of the interventions and how this might impact potential scale-up.

Opportunities for Future Research

Evidence suggests the spread, scale-up, and sustainability of health care innovations are influenced by a broad range of factors related to the people who receive (ie, participants) and deliver (ie, health care practitioners and providers) the innovations as well as by numerous organizational and societal factors [55,56,74]. The experimental studies in our review focused on a relatively narrow range of outcomes, omitted the individual and organizational provider stakeholder levels, and could not replicate the complexity of implementation in routine clinical practice. Robust experimental evaluation of effectiveness and safety is critical before real-world implementation [78]. However, it is now critical to evaluate the implementation of

proven interventions in routine clinical practice, preferably as a complementary adjunct to existing center-based programs, to incorporate the critical element of consumer choice that we and others have advocated [45,73]. Such studies should target all key implementation constructs across all relevant stakeholder levels, be embedded fully in routine clinical practice, and be evaluated for a sufficient duration to enable comprehensive assessment of all factors that contribute to successful, scalable, and sustainable implementation [74].

While it was beyond the scope of this review, an understanding of the relative importance of different factors on the implementation of telerehabilitation interventions in clinical practice is also needed. Although effective implementation is understood to be an interactive combination of factors [74], our review highlights the variability in the assessment and reporting of implementation constructs. It is unknown if this variability was due to, for example, researchers' perceived importance of specific factors when selecting the study outcomes, evidence for the differential impact of implementation factors on outcomes, or the feasibility of evaluating multiple factors within a trial design. Improving the translation of interventions into routine clinical practice requires a greater understanding of the roles of implementation factors and consistent measurement of their impact. Research that explores the relative importance of such factors would greatly advance our ability to effectively scale interventions.

Strengths and Limitations

A major strength of this review was the use of a robust systematic review methodology to understand a novel research area. Secondly, while broad variability of intervention designs across a small number of studies makes it difficult to determine how to optimize telerehabilitation for translation into clinical practice, it provides a broad overview of potential issues that could be associated with implementation and scale-up.

The findings of our review are primarily limited by a lack of studies that have evaluated cardiac telerehabilitation interventions when implemented in routine clinical practice. The assessment of implementation-related outcomes during controlled experimental studies provides some insight, but marked differences within the context of real-world rehabilitation service delivery limit their generalizability. Additionally, while promising early evidence for effectiveness, safety, acceptability, and cost of cardiac telerehabilitation interventions [33-37] suggest they could play an important role in increasing overall participation in CR, it remains unclear whether positive trial outcomes will be retained following translation into clinical practice [76]. Our review includes a small number of studies with relatively small sample sizes and homogenous cohorts. This may limit generalizability to population subgroups who are typically underserved by CR, including older adults, women, people living in regional or rural areas, and people with diverse non-English-speaking cultural backgrounds [41-45,60,64]. Finally, there was a lack of studies from developing countries where telerehabilitation could have an even greater impact due to a low provision of traditional center-based CR [79-81].

Conclusions

Cardiac telerehabilitation interventions appear to be acceptable and appropriate for many participants in experimental trials and may be a cost-effective way to increase the reach and utilization of CR. However, explicit implementation studies are urgently needed to inform best-practice translation into routine clinical practice. When possible, such studies should implement telerehabilitation in parallel with existing center-based programs so consumers can autonomously match program delivery models to their individual needs and preferences.

Acknowledgments

We would like to thank Ms Rachel West, Liaison Librarian, at Melbourne Burwood Campus, Deakin University, for her guidance in finalizing the search strategy. NS received a Deakin University Postgraduate Research Scholarship for his PhD from the university.

Authors' Contributions

NS, JR, and RM designed the study. NS conducted the search; removed the duplicates; screened the titles, abstracts, and full text of the studies; and drafted the systematic review manuscript. JR and HK contributed to verifying screening, and JR contributed to the full-text review and drafting the manuscript. All coauthors—JR, LG, HK, and RM—contributed to the critical revision of the manuscript and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies for included studies.
[DOCX File , 23 KB - mhealth_v8i11e17957_app1.docx]

Multimedia Appendix 2
Characteristics of included studies.

[DOCX File , 37 KB - mhealth_v8i11e17957_app2.docx]

References

1. Cardiovascular diseases. World Health Organization. URL: https://www.who.int/cardiovascular_diseases/en/ [accessed 2019-02-07]
2. Disease expenditure in Australia. Australian Institute of Health and Welfare. 2019 Jun 13. URL: https://www.aihw.gov.au/reports/health-welfare-expenditure/disease-expenditure-australia/contents/summary [accessed 2019-02-07]
3. Bloom DE, Cafiero ET, Jané-Llopis E, Abrahams-Gessel S, Bloom LR, Fathima S, et al. The Global Economic Burden of Noncommunicable Diseases. Geneva, Switzerland: World Economic Forum; 2011. URL: http://www3.weforum.org/docs/WEF_Harvard_HE_GlobalEconomicBurdenNonCommunicableDiseases_2011.pdf [accessed 2020-11-12]
4. Benjamin E, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2019 Update: A report from the American Heart Association. Circulation 2019 Mar 05;139(10):e56-e528. [doi: 10.1161/CIR.0000000000000659] [Medline: 30700139]
5. Coronary heart disease. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/reports/australias-health/coronary-heart-disease [accessed 2019-02-07]
6. GBD 2017 DALYsHALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1859-1922 [FREE Full text] [doi: 10.1016/S0140-6736(18)32335-3] [Medline: 30415748]
7. Balady GJ, Williams MA, Ades PA, Bittner V, Comoss P, Foody JM, et al. Core components of cardiac rehabilitation/secondary prevention programs: 2007 update. Circulation 2007 May 22;115(20):2675-2682. [doi: 10.1161/circulationaha.106.180945]
8. Cowie A, Buckley J, Doherty P, Furze G, Hayward J, Hinton S, British Association for Cardiovascular PreventionRehabilitation (BACPR). Standards and core components for cardiovascular disease prevention and rehabilitation. Heart 2019 Apr;105(7):510-515 [FREE Full text] [doi: 10.1136/heartjnl-2018-314206] [Medline: 30700518]
9. Writing Group Members, Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, American Heart Association Statistics Committee, Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2016 Update: A report from the American Heart Association. Circulation 2016 Jan 26;133(4):e38-e360. [doi: 10.1161/CIR.0000000000000350] [Medline: 26673558]
10. Piepoli MF, Corrà U, Adamopoulos S, Benzer W, Bjarnason-Wehrens B, Cupples M, et al. Secondary prevention in the clinical management of patients with cardiovascular diseases. Core components, standards and outcome measures for referral and delivery: A Policy Statement from the Cardiac Rehabilitation Section of the European Association for Cardiovascular Prevention & Rehabilitation. Endorsed by the Committee for Practice Guidelines of the European Society of Cardiology. Eur J Prev Cardiol 2014 Jun;21(6):664-681. [doi: 10.1177/2047487312449597] [Medline: 22718797]
11. Woodruffe S, Neubeck L, Clark RA, Gray K, Ferry C, Finan J, et al. Australian Cardiovascular Health and Rehabilitation Association (ACRA) core components of cardiovascular disease secondary prevention and cardiac rehabilitation 2014. Heart Lung Circ 2015 May;24(5):430-441. [doi: 10.1016/j.hlc.2014.12.008] [Medline: 25637253]
12. Taylor RS, Brown A, Ebrahim S, Jolliffe J, Noorani H, Rees K, et al. Exercise-based rehabilitation for patients with coronary heart disease: Systematic review and meta-analysis of randomized controlled trials. Am J Med 2004 May 15;116(10):682-692. [doi: 10.1016/j.amjmed.2004.01.009] [Medline: 15121495]
13. Wenger N, Froelicher E, Smith L, Ades P, Berra K, Blumenthal J, et al. Cardiac rehabilitation as secondary prevention. Agency for Health Care Policy and Research and National Heart, Lung, and Blood Institute. Clin Pract Guidel Quick Ref Guide Clin 1995 Oct(17):1-23. [Medline: 8595435]
14. Anderson L, Taylor R. Cardiac rehabilitation for people with heart disease: An overview of Cochrane systematic reviews. Int J Cardiol 2014 Dec 15;177(2):348-361. [doi: 10.1016/j.ijcard.2014.10.011] [Medline: 25456575]
15. Shields GE, Wells A, Doherty P, Heagerty A, Buck D, Davies LM. Cost-effectiveness of cardiac rehabilitation: A systematic review. Heart 2018 Sep;104(17):1403-1410 [FREE Full text] [doi: 10.1136/heartjnl-2017-312809] [Medline: 29654096]
16. Wong WP, Feng J, Pwee KH, Lim J. A systematic review of economic evaluations of cardiac rehabilitation. BMC Health Serv Res 2012 Aug 08;12:243 [FREE Full text] [doi: 10.1186/1472-6963-12-243] [Medline: 22873828]
17. De Gruyter E, Ford G, Stavreski B. Economic and social impact of increasing uptake of cardiac rehabilitation services--A cost benefit analysis. Heart Lung Circ 2016 Feb;25(2):175-183. [doi: 10.1016/j.hlc.2015.08.007] [Medline: 26442971]
18. Anderson L, Oldridge N, Thompson D, Zwisler A, Rees K, Martin N, et al. Exercise-based cardiac rehabilitation for coronary heart disease: Cochrane systematic review and meta-analysis. J Am Coll Cardiol 2016 Jan 05;67(1):1-12 [FREE Full text] [doi: 10.1016/j.jacc.2015.10.044] [Medline: 26764059]
19. Kotseva K, Wood D, De Bacquer D, EUROASPIRE Investigators. Determinants of participation and risk factor control according to attendance in cardiac rehabilitation programmes in coronary patients in Europe: EUROASPIRE IV survey. Eur J Prev Cardiol 2018 Aug;25(12):1242-1251. [doi: 10.1177/2047487318781359] [Medline: 29873511]
20. Scott IA, Lindsay KA, Harden HE. Utilisation of outpatient cardiac rehabilitation in Queensland. Med J Aust 2003 Oct 06;179(7):341-345. [doi: 10.5694/j.1326-5377.2003.tb05588.x]

21. Squires RW. Are cardiac rehabilitation programs underutilized by patients with coronary heart disease? Nat Clin Pract Cardiovasc Med 2008 Apr;5(4):192-193. [doi: 10.1038/ncpcardio1128] [Medline: 18268514]
22. Clark AM, King-Shier KM, Thompson DR, Spaling MA, Duncan AS, Stone JA, et al. A qualitative systematic review of influences on attendance at cardiac rehabilitation programs after referral. Am Heart J 2012 Dec;164(6):835-845.e2. [doi: 10.1016/j.ahj.2012.08.020] [Medline: 23194483]
23. Dunlay SM, Witt BJ, Allison TG, Hayes SN, Weston SA, Koepsell E, et al. Barriers to participation in cardiac rehabilitation. Am Heart J 2009 Nov;158(5):852-859 [FREE Full text] [doi: 10.1016/j.ahj.2009.08.010] [Medline: 19853708]
24. Rose M, Timmons SM, Amerson R, Reimels E, Pruitt RH. Facilitators and barriers in cardiac rehabilitation participation: An integrative review. J Nurse Pract 2011 May;7(5):399-408. [doi: 10.1016/j.nurpra.2011.02.003]
25. Ruano-Ravina A, Pena-Gil C, Abu-Assi E, Raposeiras S, van 't Hof A, Meindersma E, et al. Participation and adherence to cardiac rehabilitation programs. A systematic review. Int J Cardiol 2016 Nov 15;223:436-443. [doi: 10.1016/j.ijcard.2016.08.120] [Medline: 27557484]
26. Neubeck L, Freedman SB, Clark AM, Briffa T, Bauman A, Redfern J. Participating in cardiac rehabilitation: A systematic review and meta-synthesis of qualitative data. Eur J Prev Cardiol 2012 Jun;19(3):494-503. [doi: 10.1177/1741826711409326] [Medline: 22779092]
27. Ades P, Waldmann M, McCann W, Weaver S. Predictors of cardiac rehabilitation participation in older coronary patients. Arch Intern Med 1992 May;152(5):1033-1035. [Medline: 1580707]
28. Ferguson E. Cardiac rehabilitation: An effective and comprehensive but underutilized program to reduce cardiovascular risk in patients with CVD. US Cardiol Rev 2006;2:14e6. [doi: 10.15420/usc.2006.3.2.14]
29. Dalal H, Evans P. Achieving national service framework standards for cardiac rehabilitation and secondary prevention. BMJ 2003 Mar 01;326(7387):481-484 [FREE Full text] [doi: 10.1136/bmj.326.7387.481] [Medline: 12609946]
30. Dalal HM, Zawada A, Jolly K, Moxham T, Taylor RS. Home based versus centre based cardiac rehabilitation: Cochrane systematic review and meta-analysis. BMJ 2010 Jan 19;340:b5631 [FREE Full text] [doi: 10.1136/bmj.b5631] [Medline: 20085991]
31. Jackson AC, Higgins RO, Murphy BM, Rogerson M, Le Grande MR. Cardiac rehabilitation in Australia: A brief survey of program characteristics. Heart Lung Circ 2018 Dec;27(12):1415-1420. [doi: 10.1016/j.hlc.2017.08.024] [Medline: 29100840]
32. Parmanto B, Saptono A. Telerehabilitation: State-of-the-Art from an Informatics Perspective. Int J Telerehabil 2009 Sep;1(1):73-84 [FREE Full text] [doi: 10.5195/ijt.2009.6015] [Medline: 25945164]
33. 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;22(1):35-74. [doi: 10.1177/2047487313501093] [Medline: 23943649]
34. Frederix I, Vanhees L, Dendale P, Goetschalckx K. A review of telerehabilitation for cardiac patients. J Telemed Telecare 2014 Dec 04;21(1):45-53. [doi: 10.1177/1357633x14562732]
35. Huang K, Liu W, He D, Huang B, Xiao D, Peng Y, et al. Telehealth interventions versus center-based cardiac rehabilitation of coronary artery disease: A systematic review and meta-analysis. Eur J Prev Cardiol 2015 Aug;22(8):959-971. [doi: 10.1177/2047487314561168] [Medline: 25488550]
36. Jin K, Khonsari S, Gallagher R, Gallagher P, Clark AM, Freedman B, et al. Telehealth interventions for the secondary prevention of coronary heart disease: A systematic review and meta-analysis. Eur J Cardiovasc Nurs 2019 Apr;18(4):260-271. [doi: 10.1177/1474515119826510] [Medline: 30667278]
37. Rawstorn JC, Gant N, Direito A, Beckmann C, Maddison R. Telehealth exercise-based cardiac rehabilitation: A systematic review and meta-analysis. Heart 2016 Aug 01;102(15):1183-1192. [doi: 10.1136/heartjnl-2015-308966] [Medline: 26936337]
38. Neubeck L, Redfern JU, Fernandez R, Briffa T, Bauman A, Freedman SB. Telehealth interventions for the secondary prevention of coronary heart disease: A systematic review. Eur J Cardiovasc Prev Rehabil 2009 Jun;16(3):281-289. [doi: 10.1097/hjr.0b013e32832a4e7a]
39. Mackay MM. Australian Mobile Phone Lifestyle Index, 9th edition. Sydney, Australia: AIMIA; 2013 Oct. URL: https://www.sponsor-ed.com.au/app/webroot/uploaded_files/media/Ampli%202013%20Report_Final_October%2024.pdf [accessed 2020-11-09]
40. Neubeck L, Lowres N, Benjamin EJ, Freedman SB, Coorey G, Redfern J. The mobile revolution--Using smartphone apps to prevent cardiovascular disease. Nat Rev Cardiol 2015 Jun;12(6):350-360. [doi: 10.1038/nrcardio.2015.34] [Medline: 25801714]
41. 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. JAMA 2015;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]
42. Pfaeffli Dale L, 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]
43. 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]

44. Varnfield M, Karunanithi M, Lee C, Honeyman E, Arnold D, Ding H, et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: Results from a randomised controlled trial. Heart 2014 Nov;100(22):1770-1779 [FREE Full text] [doi: 10.1136/heartjnl-2014-305783] [Medline: 24973083]
45. Maddison R, Rawstorn JC, Stewart RAH, Benatar J, Whittaker R, Rolleston A, et al. Effects and costs of real-time cardiac telerehabilitation: Randomised controlled non-inferiority trial. Heart 2019 Jan;105(2):122-129 [FREE Full text] [doi: 10.1136/heartjnl-2018-313189] [Medline: 30150328]
46. Milat AJ, Bauman AE, Redman S, Curac N. Public health research outputs from efficacy to dissemination: A bibliometric analysis. BMC Public Health 2011 Dec 15;11:934 [FREE Full text] [doi: 10.1186/1471-2458-11-934] [Medline: 22168312]
47. Wiltsey Stirman S, Kimberly J, Cook N, Calloway A, Castro F, Charns M. The sustainability of new programs and innovations: A review of the empirical literature and recommendations for future research. Implement Sci 2012 Mar 14;7:17 [FREE Full text] [doi: 10.1186/1748-5908-7-17] [Medline: 22417162]
48. Wolfenden L, Chai LK, Jones J, McFadyen T, Hodder R, Kingsland M, et al. What happens once a program has been implemented? A call for research investigating strategies to enhance public health program sustainability. Aust N Z J Public Health 2019 Feb;43(1):3-4. [doi: 10.1111/1753-6405.12867] [Medline: 30690829]
49. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: Models for dissemination and implementation research. Am J Prev Med 2012 Sep;43(3):337-350 [FREE Full text] [doi: 10.1016/j.amepre.2012.05.024] [Medline: 22898128]
50. Koorts H, Eakin E, Estabrooks P, Timperio A, Salmon J, Bauman A. Implementation and scale up of population physical activity interventions for clinical and community settings: The PRACTIS guide. Int J Behav Nutr Phys Act 2018 Jun 08;15(1):51 [FREE Full text] [doi: 10.1186/s12966-018-0678-0] [Medline: 29884236]
51. Reis RS, Salvo D, Ogilvie D, Lambert EV, Goenka S, Brownson RC. Scaling up physical activity interventions worldwide: Stepping up to larger and smarter approaches to get people moving. Lancet 2016 Sep;388(10051):1337-1348. [doi: 10.1016/s0140-6736(16)30728-0]
52. PROSPERO: International prospective register of systematic reviews. National Institute for Health Research. URL: https://www.crd.york.ac.uk/prospero/ [accessed 2019-02-07]
53. 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]
54. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
55. 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]
56. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
57. Rayyan QCRI: Qatar Computing Research Institute. URL: https://rayyan.qcri.org/welcome
58. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Hoboken, NJ: John Wiley & Sons; 2008.
59. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R. Systematic reviews of etiology and risk. In: JBI Manual for Evidence Synthesis. Adelaide, Australia: JBI; 2020:255-257.
60. Kraal JJ, Van den Akker-Van Marle ME, Abu-Hanna A, Stut W, Peek N, Kemps HM. Clinical and cost-effectiveness of home-based cardiac rehabilitation compared to conventional, centre-based cardiac rehabilitation: Results of the FIT@Home study. Eur J Prev Cardiol 2017 Aug;24(12):1260-1273 [FREE Full text] [doi: 10.1177/2047487317710803] [Medline: 28534417]
61. Lear S, Singer J, Banner-Lukaris D, Horvat D, Park J, Bates J, et al. Improving access to cardiac rehabilitation using the internet: A randomized trial. Stud Health Technol Inform 2015;209:58-66. [Medline: 25980706]
62. Park LG, Howie-Esquivel J, Chung ML, Dracup K. A text messaging intervention to promote medication adherence for patients with coronary heart disease: A randomized controlled trial. Patient Educ Couns 2014 Feb;94(2):261-268. [doi: 10.1016/j.pec.2013.10.027] [Medline: 24321403]
63. Piotrowicz E, Korzeniowska-Kubacka I, Chrapowicka A, Wolszakiewicz J, Dobraszkiewicz-Wasilewska B, Batogowski M, et al. Feasibility of home-based cardiac telerehabilitation: Results of TeleInterMed study. Cardiol J 2014 Oct 29;21(5):539-546. [doi: 10.5603/cj.a2014.0005]
64. Salvi D, Ottaviano M, Muuraiskangas S, Martínez-Romero A, Vera-Muñoz C, Triantafyllidis A, et al. An m-Health system for education and motivation in cardiac rehabilitation: The experience of HeartCycle guided exercise. J Telemed Telecare 2017 Mar 28;24(4):303-316. [doi: 10.1177/1357633x17697501]

65. Redfern J, Santo K, Coorey G, Thakkar J, Hackett M, Thiagalingam A, et al. Factors influencing engagement, perceived usefulness and behavioral mechanisms associated with a text message support program. PLoS One 2016;11(10):e0163929 [FREE Full text] [doi: 10.1371/journal.pone.0163929] [Medline: 27741244]
66. Kraal JJ, Peek N, van den Akker-Van Marle ME, Kemps HM. Effects and costs of home-based training with telemonitoring guidance in low to moderate risk patients entering cardiac rehabilitation: The FIT@Home study. BMC Cardiovasc Disord 2013 Oct 08;13:82 [FREE Full text] [doi: 10.1186/1471-2261-13-82] [Medline: 24103384]
67. Banner D, Lear S, Kandola D, Singer J, Horvat D, Bates J, et al. The experiences of patients undertaking a 'virtual' cardiac rehabilitation program. Stud Health Technol Inform 2015;209:9-14. [Medline: 25980699]
68. Pfaeffli Dale L, Whittaker R, Dixon R, Stewart R, Jiang Y, Carter K, et al. Acceptability of a mobile health exercise-based cardiac rehabilitation intervention. J Cardiopulm Rehabil Prev 2015;35(5):312-319. [doi: 10.1097/hcr.0000000000000125]
69. Rawstorn JC, Gant N, Rolleston A, Whittaker R, Stewart R, Benatar J, et al. End users want alternative intervention delivery models: Usability and acceptability of the REMOTE-CR exercise-based cardiac telerehabilitation program. Arch Phys Med Rehabil 2018 Nov;99(11):2373-2377. [doi: 10.1016/j.apmr.2018.06.027] [Medline: 30076800]
70. Szalewska D, Zieliński P, Tomaszewski J, Kusiak-Kaczmarek M, Łepska L, Gierat-Haponiuk K, et al. Effects of outpatient followed by home-based telemonitored cardiac rehabilitation in patients with coronary artery disease. Kardiol Pol 2015;73(11):1101-1107. [doi: 10.5603/KP.a2015.0095] [Medline: 25987401]
71. Skobel E, Knackstedt C, Martinez-Romero A, Salvi D, Vera-Munoz C, Napp A, et al. Internet-based training of coronary artery patients: The Heart Cycle Trial. Heart Vessels 2017 Apr;32(4):408-418. [doi: 10.1007/s00380-016-0897-8] [Medline: 27730298]
72. Wandersman A, Duffy J, Flaspohler P, Noonan R, Lubell K, Stillman L, et al. Bridging the gap between prevention research and practice: The interactive systems framework for dissemination and implementation. Am J Community Psychol 2008 Jun;41(3-4):171-181. [doi: 10.1007/s10464-008-9174-z] [Medline: 18302018]
73. Dalal H, Doherty P, Taylor R. Cardiac rehabilitation. BMJ 2015 Sep 29;351:h5000 [FREE Full text] [doi: 10.1136/bmj.h5000] [Medline: 26419744]
74. Chaudoir SR, Dugan AG, Barr CHI. Measuring factors affecting implementation of health innovations: A systematic review of structural, organizational, provider, patient, and innovation level measures. Implement Sci 2013 Feb 17;8:22 [FREE Full text] [doi: 10.1186/1748-5908-8-22] [Medline: 23414420]
75. Karsh B. Beyond usability: Designing effective technology implementation systems to promote patient safety. Qual Saf Health Care 2004 Oct 01;13(5):388-394. [doi: 10.1136/qshc.2004.010322]
76. Brownson R, Colditz G, Proctor E, editors. Dissemination and Implementation Research in Health: Translating Science to Practice. 2nd edition. New York, NY: Oxford University Press; 2018.
77. Greenhalgh T, A'Court C, Shaw S. Understanding heart failure; explaining telehealth - A hermeneutic systematic review. BMC Cardiovasc Disord 2017 Jun 14;17(1):156 [FREE Full text] [doi: 10.1186/s12872-017-0594-2] [Medline: 28615004]
78. Dearing JW. Evolution of diffusion and dissemination theory. J Public Health Manag Pract 2008;14(2):99-108. [doi: 10.1097/01.phh.0000311886.98627.b7]
79. Grace SL, Turk-Adawi KI, Contractor A, Atrey A, Campbell N, Derman W, et al. Cardiac rehabilitation delivery model for low-resource settings. Heart 2016 Sep 15;102(18):1449-1455 [FREE Full text] [doi: 10.1136/heartjnl-2015-309209] [Medline: 27181874]
80. Korenfeld Y, Mendoza-Bastidas C, Saavedra L, Montero-Gómez A, Perez-Terzic C, Thomas RJ, et al. Current status of cardiac rehabilitation in Latin America and the Caribbean. Am Heart J 2009 Sep;158(3):480-487. [doi: 10.1016/j.ahj.2009.06.020] [Medline: 19699874]
81. Oldridge NB, Pakosh MT, Thomas RJ. Cardiac rehabilitation in low- and middle-income countries: A review on cost and cost-effectiveness. Int Health 2016 Mar;8(2):77-82. [doi: 10.1093/inthealth/ihv047] [Medline: 26208507]

Abbreviations

CHD: coronary heart disease
CR: cardiac rehabilitation
CVD: cardiovascular disease
ICT: information and communication technology
MI: myocardial infarction
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial

Edited by C Dias; submitted 23.01.20; peer-reviewed by SY Chen, S Olsen, G Chappell; comments to author 17.03.20; revised version received 02.07.20; accepted 15.07.20; published 27.11.20.

Please cite as:
Subedi N, Rawstorn JC, Gao L, Koorts H, Maddison R
Implementation of Telerehabilitation Interventions for the Self-Management of Cardiovascular Disease: Systematic Review
JMIR Mhealth Uhealth 2020;8(11):e17957
URL: http://mhealth.jmir.org/2020/11/e17957/
doi:10.2196/17957
PMID:33245286

Original Paper

Exploitation of Outgoing and Incoming Telephone Calls in the Context of Circadian Rhythms of Social Activity Among Elderly People: Observational Descriptive Study

Timothée Aubourg[1,2,3], MSc; Jacques Demongeot[2,3,4], MD, PhD; Hervé Provost[1,3], MSc; Nicolas Vuillerme[2,3,4], PhD

[1]Orange Labs, Meylan, France

[2]Univ Grenoble Alpes, AGEIS, Grenoble, France

[3]LabCom Telecom4Health, Univ Grenoble Alpes & Orange Labs, Grenoble, France

[4]Institut Universitaire de France, Paris, France

Corresponding Author:
Timothée Aubourg, MSc
Orange Labs
Chemin du Vieux Chêne
Meylan
France
Phone: 33 0950399591
Email: timotheeaubourg@gmail.com

Abstract

Background: In the elderly population, analysis of the circadian rhythms of social activity may help in supervising homebound disabled and chronically ill populations. Circadian rhythms are monitored over time to determine, for example, the stability of the organization of daily social activity rhythms and the occurrence of particular desynchronizations in the way older adults act and react socially during the day. Recently, analysis of telephone call detail records has led to the possibility of determining circadian rhythms of social activity in an objective unobtrusive way for young patients from their outgoing telephone calls. At this stage, however, the analysis of incoming call rhythms and the comparison of their organization with respect to outgoing calls remains to be performed in underinvestigated populations (in particular, older populations).

Objective: This study investigated the persistence and synchronization of circadian rhythms in telephone communication by older adults.

Methods: The study used a longitudinal 12-month data set combining call detail records and questionnaire data from 26 volunteers aged 70 years or more to determine the existence of persistent and synchronized circadian rhythms in their telephone communications. The study worked with the following four specific telecommunication parameters: (1) recipient of the telephone call (alter), (2) time at which the call began, (3) duration of the call, and (4) direction of the call. We focused on the following two issues: (1) the existence of persistent circadian rhythms of outgoing and incoming telephone calls in the older population and (2) synchronization with circadian rhythms in the way the older population places and responds to telephone calls.

Results: The results showed that older adults have their own specific circadian rhythms for placing telephone calls and receiving telephone calls. These rhythms are partly structured by the way in which older adults allocate their communication time over the day. In addition, despite minor differences between circadian rhythms for outgoing and incoming calls, our analysis suggests the two rhythms could be synchronized.

Conclusions: These results suggest the existence of potential persistent and synchronized circadian rhythms in the outgoing and incoming telephone activities of older adults.

(JMIR Mhealth Uhealth 2020;8(11):e13535) doi:10.2196/13535

KEYWORDS

circadian rhythms; phone call detail records; older population; digital phenotype

Introduction

Background

The lives of humans are deeply structured on a daily basis, which both reflects and contributes to the emergence of regular rhythms of activity, the so-called *circadian* (ie, 24 hours) rhythms [1]. Circadian rhythms help individuals affirm and maintain stability in their lives. Behind the apparent regularity, circadian rhythms reflect complex processes with multiple origins. On one hand, at the endogenous level, they stem directly from a biological clock, known as the suprachiasmatic nucleus [2]. Depending on the individual, this internal clock has an almost 24-hour period and regulates the functioning of the entire body throughout the day. On the other hand, at the exogenous level, these biological time posts are entangled with physical and social cues [3], such as the day-night alternation [4] and the daily social interactions [5] that occur between individuals and their social network. These external "time givers," known under the German word "Zeitgeberen," permit individuals to adapt themselves to their physical and social environments by aligning their daily activities to a precise and persistent 24-hour clock [4]. Overall, this close synchronization results from the continuous interplay between the fundamental biological and social time cycles in the lives of all individuals [6].

According to the social zeitgeber theory [7], a change in social cues or the occurrence of particular social events, such as life stress [7] and affective episodes [8], may have severe repercussions on these biological and social times. Notably, social changes have the potential to weaken the stability of individuals' daily lives [7], and the resulting instability may be associated with the social jet-lag phenomenon, in which the biological and social circadian rhythms are desynchronized [9]. In the field of health monitoring, such desynchronization provides important information about an individual's health status because it may be directly associated with health complications, such as sleep disruption [10], depression [11], and cardiometabolic problems [12]. For health professionals, it is important to properly manage these social disruptions. Thus, their detection may permit us to better understand the mechanisms underlying these severe health issues and thereby prevent them.

For seniors in particular, monitoring circadian rhythms is of importance, as amply demonstrated by the rich literature on the subject, in which these rhythms were shown to change greatly with age (eg, [13-16]). Such changes are generally due to a degradation in temporal organization in the life of seniors caused by age-related alterations that affect their biological clock [15]. These alterations are reflected, for instance, by the chronotype shift that usually appears with aging [14]. This behavioral drift can be considered as a time-induced transition in a senior's preferences for being active during the day, generally in the morning [17]. At the social level of seniors, this chronotype shift may notably be associated with alterations in their social routines, such as those imposed by social-rhythm stereotyping [18]. Moreover, the chronotype shift may be accompanied by the nonreception of social cues caused by the aging process, leading to a decline in social interactions [19,20], inconsistent meal times [21,22], or fragmented sleep [17]. Finally, these social disruptions may cause feedback and impact the biological state of the individual [7]. Such feedback may notably maintain or worsen the dysregulation of biological circadian rhythms, thus facilitating the health issues mentioned above, which may dramatically impact an older person's life, as evinced by Alzheimer disease [23-25], Parkinson disease [26], or dementia [21]. Thus, it is of importance to investigate the social aspects of circadian rhythms and their disruption in the elderly population to better understand how these rhythms, or lack thereof, affect the biological clock of seniors and the mechanisms that trigger health problems.

Prior Studies

A body of literature is presently emerging around mobile health sensing, including the effect of circadian rhythms and social behaviors (eg, [27-29]). Much of this literature discusses the potential of using telephones to enhance the traditional health care system by exploiting health-related data generated by telephone use [30,31]. In particular, analyses of telecommunication data, such as telephone call detail records, enable the detection of circadian rhythms of social activity [32-36]. For example, the recent work of Aledavood et al [32] demonstrated the existence of a circadian temporal signature of outgoing telephone call activity in a young population. Additionally, with the same population sample, Aledavood et al [33] showed that when using a descriptive approach to compare outgoing telephone calls with outgoing text messages, young adults establish persistent temporal signatures whose patterns may differ between voice calls and text messages. In the field of health care research, these fundamental results about digital rhythms suggest that telephones may be used to monitor daily social activity. In recent publications, this sort of exploitation of telephones has provided promising results for monitoring the health status and biological timeframe of individuals [37-40].

Note, however, that it remains unclear whether similar results apply to underinvestigated populations and how they can be implemented. In particular, no studies have heretofore investigated whether the telephone activity of seniors follows a consistent and synchronized circadian rhythm.

Study Goals

This work was thus specifically designed to determine whether persistent and synchronized circadian rhythms exist in the telephone activity of older adults. To this end, we used a longitudinal data set of 12 successive months combining call detail records and questionnaire data from 21 volunteers over 70 years of age. In this study, we extracted the following four specific telecommunication parameters from the call detail records: (1) recipient of the telephone call (alter), (2) time at which the telephone call began, (3) telephone call duration, and (4) telephone call direction (outgoing or incoming).

Methods

Data Collection and Volunteer Recruitment

This study involved a secondary analysis of previously published and unpublished data set analyses [41-43]. Our data set included

12 months of outgoing call detail records for 26 volunteers (20 women and 6 men; median age 84 years, range 71-91 years). Call detail records provided by the local communication service provider were collected from the personal telephones of the volunteers. Each call detail record contained the date, hour, source ID, recipient ID, direction, and duration of the call (in seconds). In addition, volunteers with several telephones registered with their communication service provider (eg, one or more landlines and/or one or more mobile telephones) provided call detail records for all their telephones. Note that the telephone owners and the telephone contacts remained anonymous. Additionally, each volunteer filled in a questionnaire about the people in their telephone social network and classified each of their telephone contacts into one of the following five social categories: family, friend, acquaintance, health professional, and other.

This study and its corresponding experimental protocols were declared to the French Data Protection Authority (CNIL registered data protection officer, France Telecom 2011 n°44). All experimental methods were carried out in accordance with the relevant regulations, and written informed consent was obtained from all participants before data were collected and anonymized.

Data Preprocessing

Participants did not all enroll in the survey at the same time, so the dates of inclusion varied. Thus, the call detail record data set was filtered to select the time interval when the greatest number of volunteers were actively participating in the study. The call detail record data set was then preprocessed by applying the method described by Saramäki et al [44], which involved selecting only those participants who used their telephones throughout the 12-month observation period, providing a set of 21 individuals (Table 1). Additionally, note that only the incoming telephone calls answered by the volunteer were used in the analysis (ie, incoming telephone calls that went unanswered were removed from the call detail record data set).

Table 1. Structure of the data set of call detail records before and after preprocessing.

Variable	Before preprocessing	After preprocessing
Number of participants	26	21
Female	20	16
Male	6	5
Age (years), range	71-91	71-91
Age (years), mean (SD)	84 (4)	83 (4)
Total number of calls		
Outgoing	19,198	18,338
Incoming	20,062	17,879
Average number of calls per individual, first quartile		
Outgoing	285	481
Incoming	579	605
Average number of calls per individual, median		
Outgoing	590	710
Incoming	718	800
Average number of calls per individual, third quartile		
Outgoing	944	1096
Incoming	947	992

Data Analysis

Measuring the Daily Rhythms of Telephone Activity

We followed the descriptive approach proposed by Aledavood et al [32] that treats outgoing and incoming telephone calls separately. This approach involves calculating the daily pattern of telephone activity for the study's entire 12-month data set, which is done by applying a two-step process consisting of (1) coarse graining the time dimension into unique days divided into 24 one-hour time slots and (2) calculating the average fraction of daily calls. The fraction of daily telephone calls for each time slot is given by the following formula:

where $n(t)$ is the number of calls in time slot t.

We followed this approach at the aggregate level to obtain a concise overview of the data set's structure and trends. Thereafter, to avoid an ecological fallacy [32], the daily rhythms of telephone activity for each individual (ie, ego) was estimated by zooming in to the individual level, as done by Aledavood et al [32].

Determining the Consistence of Daily Rhythms of Telephone Activity in Older Adults

To determine the consistence of daily rhythms of telephone calls for each ego, we followed an analytical approach based on the notion of persistence applied separately to outgoing and incoming telephone calls. This notion was proposed by Saramäki et al [44] and applied previously [32-34] to assess the repeatability in time of daily rhythms of telephone activity in accordance with circadian rhythm framing. The approach consists of comparing the stability of estimated daily rhythms at distinct successive time points by applying the below three steps.

(1) Step 1: *Time discretization.* We applied a coarse-graining process [32]. Since the call detail record data set contained 12 months of observations, it was split into three successive 4-month periods called T1, T2, and T3. Note that, although the literature offers no clear consensus regarding the time-interval size, aggregating data into intervals of several months enhances the robustness of the persistence analysis and limits the short-term variations in the patterns of telephone calls [45].

(2) Step 2: *Calculation of daily rhythm*. We calculated the daily rhythm of telephone calls for each ego and for each time interval (T1-T3) by using the same two-step computation described in the previous section.

(3) Step 3: *Persistence of daily rhythm.* The persistence of daily rhythm was determined for each ego by measuring the repeatability of the daily rhythms of telephone call activity over time via a persistence analysis. This analysis involves calculating the stability of each ego's call patterns in each time period (T1-T3) by using the square root of the Jensen-Shannon divergence dissimilarity measure [35] (see Statistical Tools in the Methods section).

We denote by D_{self} the dissimilarity measure of the individual's daily rhythms between two successive periods. D_{self} is given by the following formula:

where P_i^T (P_i^{T+1}) is the discrete probability distribution of the call fraction of ego i at time period T (T+1) and JSD is Jensen-Shannon divergence.

We denote by $\langle D_{self} \rangle$ the average of D_{self} as follows:

where N_T=3 is the number of time periods.

Thereafter, a reference scale for further comparison was implemented for each ego by calculating, for each time interval (T1-T3), the square root of the Jensen-Shannon divergence between the daily pattern of the given ego and that of every other ego.

We denote by D_{ref} the dissimilarity measure between two daily rhythms for two distinct individuals in the same time period. D_{ref} is given by the following formula:

where P_i^T (P_j^T) is the discrete probability distribution of call fractions for ego i (j) at time period T, with $i \neq j$, and JSD is Jensen-Shannon divergence.

We denote by $\langle D_{ref} \rangle$ the average of D_{ref}, which is given as follows:

where n=21 is the number of individuals.

Finally, the persistence of daily rhythm for a given ego was evaluated by comparing the average call pattern over time with the average reference scale. Formally, an ego's daily rhythm is persistent if and only if $\langle D_{self} \rangle / \langle D_{ref} \rangle < 1$.

Comparing Daily Rhythms of Outgoing and Incoming Telephone Activity

The daily rhythms of outgoing and incoming telephone calls are compared by using the following two steps:

(1) Step 1: *Synchronization assessment*. For each individual and for each time interval (T1-T3), we applied the Kolmogorov-Smirnov comparison test to compare the distribution of outgoing telephone calls during the day with the distribution of incoming telephone calls [46]. The two distributions are considered synchronized if and only if they follow the same law with a P value >.05 under the following null hypothesis: "H_0: the distributions of outgoing and incoming telephone calls follow the same law."

(2) Step 2: We completed the synchronization assessment by applying a descriptive approach that investigates the alter specificity of the daily rhythms of outgoing and incoming telephone calls to determine whether specific time intervals exist for communicating with specific alters during the day. To this end, we used both the call detail records and questionnaire data for the egos, as performed by Aledavood et al [32]. Based on these pieces of information, the following three questions were addressed: (1) Were there specific hours for communicating with specific alters? (2) What was the fraction of telephone calls to/from the top two alters and who were the top two alters (ie, the alters most frequently in telephone contact with the ego)? (3) What was the variation in the duration of telephone calls between egos and their social network?

In particular, to answer question 1, we analyzed how each ego corresponds with his or her alters through the day by applying the following steps (step 1 and step 2):

(1) Step 1: *Time discretization*. For each ego in each time period (T1-T3), we coarse grained the time dimension into four bins of equal time as follows: night (12 AM-6 AM), morning (6 AM-12 PM), afternoon (12 PM-6 PM), and evening (6 PM-12 AM).

(2) Step 2: *Alter-specificity assessment*. For each ego and for time periods T1 to T3, the alter specificity of the daily rhythms was assessed by comparing the alter structure with that of a null model simulating total randomness in alter specificity. To this

end, we estimated the diversity of alters for each 6-hour bin and for each ego by calculating the associated relative entropies (see Statistical Tools in the Methods section) for time periods T1 to T3. This calculation was performed in two steps as follows: (1) Step 2a: *Calculation of origin entropy*. We calculated the origin entropy H_{orig} for each ego and for each 6-hour time bin for time periods T1 to T3; (2) Step 2b: *Calculation of relative entropy*. Thereafter, we obtained the relative entropy H_{rel} by normalizing H_{orig} by the average $\langle H_{ref} \rangle$ of the reference entropy resulting from a null model. This null model is based on the hypothesis that no specific times are associated with telephone calls to specific alters. Thus, a low relative entropy, tending to zero, indicates a relevant alter specificity in the call pattern.

Statistical Tools

This section presents the following three technical tools used in this study: (1) normalization formulas, (2) Jenson-Shannon divergence, and (3) relative entropy.

Normalization

To visualize the peaks and troughs of each subject's telephone call activity on a single graph, we normalized the daily fraction of their calls to the range (0, 1) by using a unity-based normalization formula [47]. Formally, we let n=21 be the number of participants included in this study and let p_i be a length t vector of the daily fraction of calls for subject i, with t in {0, 1, …, 23} and i in {1,2, …, n}. We denote by $p_i(t)$ the ratio of calls of subject i at hour t. Then, for each element $p_i(t)$ in the vector p_i, we denote by $p_{i,\text{normalized}}(t)$ its normalized value as follows:

Jensen-Shannon Divergence

Jensen-Shannon divergence is a measure of dissimilarity that compares two probability distributions; it is a symmetrical finite-valued version of the Kullback-Leibler divergence. Its square root can be used as a metric for measuring the distance between two probability distributions. Formally, in a discrete context, we have the following:

where $p1$ and $p2$ are two discrete probability distributions, $H(\cdot)$ is the Shannon entropy, and JSD is Jensen-Shannon divergence.

Entropy Measures

At the individual level, to compare the daily rhythms of outgoing and incoming telephone calls by investigating their alter specificity, we measured the diversity of alters at different hours of the day by calculating the relative entropy. As detailed below, the relative entropy is computed separately for outgoing and incoming telephone calls.

For a given ego i, with i in {1, 2, …, n} and n=21 being the number of participants, we calculated the relative entropy of alter specificity by normalizing the origin entropy by the reference entropy for each 6-hour time interval. This calculation was performed in the following three steps:

(1) Step 1: *Origin entropy.* Formally, let $A_i = \{a_{i1}, a_{i2}, \ldots, a_{im}\}$ be the set of alters of ego i, where m gives the size of the social network of ego i. For a 6-hour time interval t, the origin entropy is as follows:

where $p_{i,k}(t)$ is the fraction of calls between ego i and the alter a_{ik} at time interval t.

(2) Step 2: *Reference entropy.* We now defined a null model that simulates a system with no specific alter structure. To this end, the number of calls between ego i and each alter is preserved, but to simulate randomness, we shuffled the time of the calls in each 2-week period [32]. Thereafter, for each 6-hour interval t for this shuffled data set, we calculated the corresponding origin entropy $H_{i,origin}(t)$ as outlined in the previous step. By iterating this process N=1000 times, we obtained the reference entropy [32].

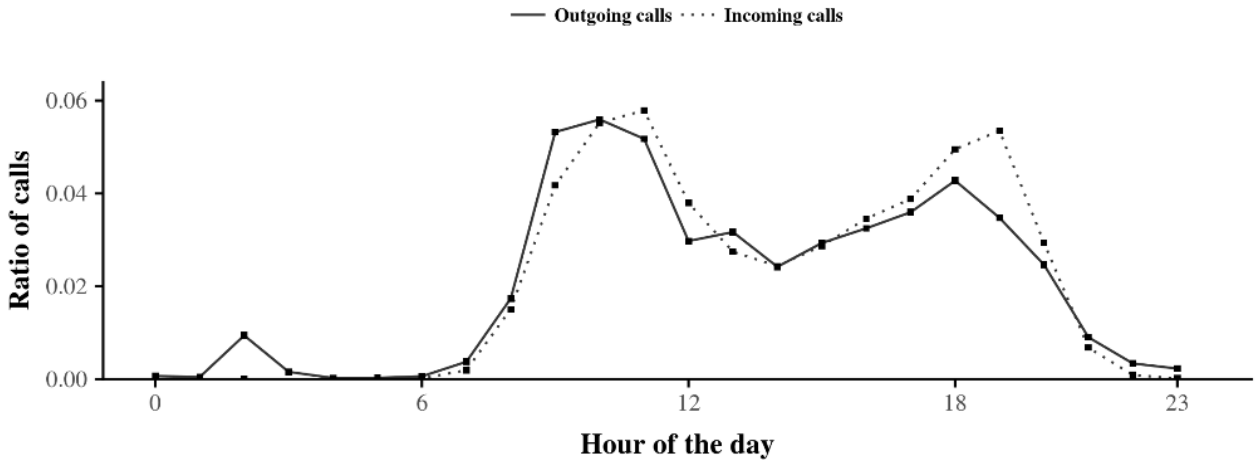
(3) Step 3: *Relative entropy.* To estimate the alter specificity of our original system, we normalized the origin entropy of ego i in each time interval t by the reference entropy as follows:

Results

Telephone Activity of Older Adults Exhibited Daily Rhythms

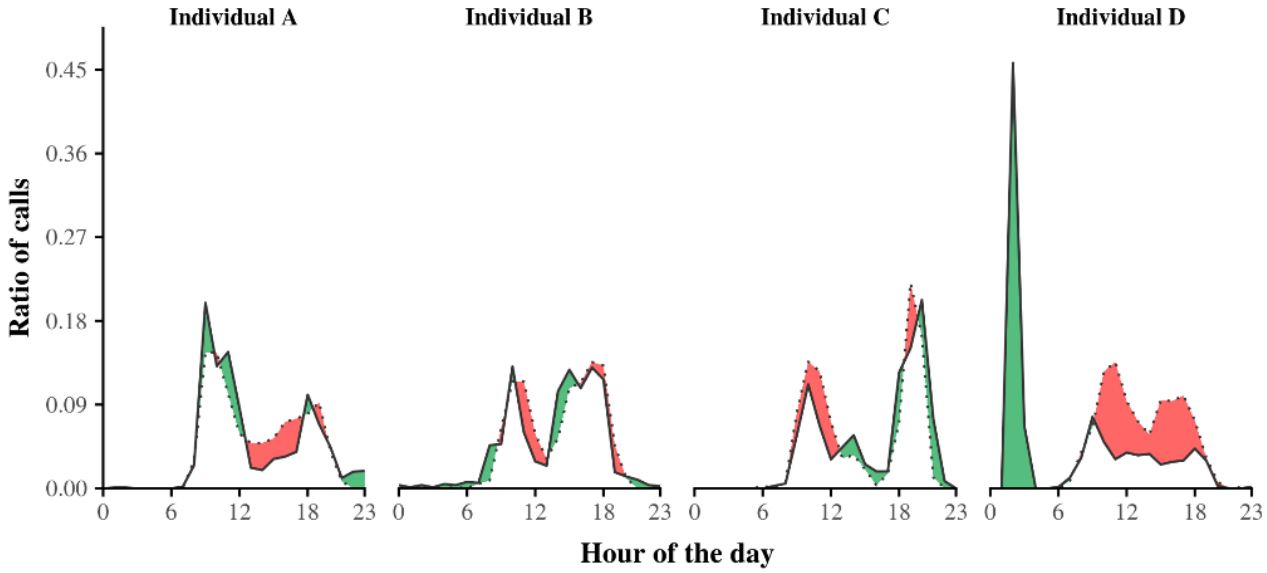
Figure 1 shows separately, at the aggregate level, the overall daily rhythms of outgoing and incoming telephone calls for the older adults participating in the study. The results showed that both of these daily rhythms were marked by the following two distinct peaks in activity: (1) a peak at mid-morning around 10 AM after typical waking times and (2) a peak at the end of the day around 6 PM. Both peaks are in accordance with the notion of circadian acrophase in humans, which is usually diurnal [2]. Note that these two peaks are separated by a period of low activity in the afternoon, beginning around 2 PM and matching the typical nap time for elderly people [48]. Additionally, telephone call activity was the slowest during the habitual sleep times in the nocturnal period, in accordance with the circadian nadir phenomena usually occurring at this time [2]. For outgoing telephone calls only, we also noted a small peak in nocturnal activity around 2 AM, which seems rather unusual given the prominent diurnal activity pattern of the data set. For incoming telephone calls, the evening activity peak was noticeably higher than that for outgoing telephone calls. Taken together, these results suggest that the daily rhythm of telephone calls in older adults corresponds with circadian rhythm framing. Interestingly, the daily rhythms for outgoing and incoming telephone calls resembled each other, although they also had their own specific characteristics.

Figure 1. Daily rhythm of telephone calls of older adults at the aggregate level for outgoing and incoming telephone calls. The shape of the curves suggests that similarities exist in how older adults place and receive telephone calls during the day. It also suggests differences, especially at night around 2 AM and in the evening when the fraction of incoming telephone calls is prominent.



To avoid an ecological fallacy [32], we also observed separately the daily telephone activity of each ego. Figure 2 shows the call fractions for four egos, for which the highest peak in telephone activity was in the morning, afternoon, evening, and night. The results suggest not only (1) the existence of daily rhythms for each ego for both outgoing and incoming telephone calls, but also (2) similarities and differences in the way the egos place and receive telephone calls during the day.

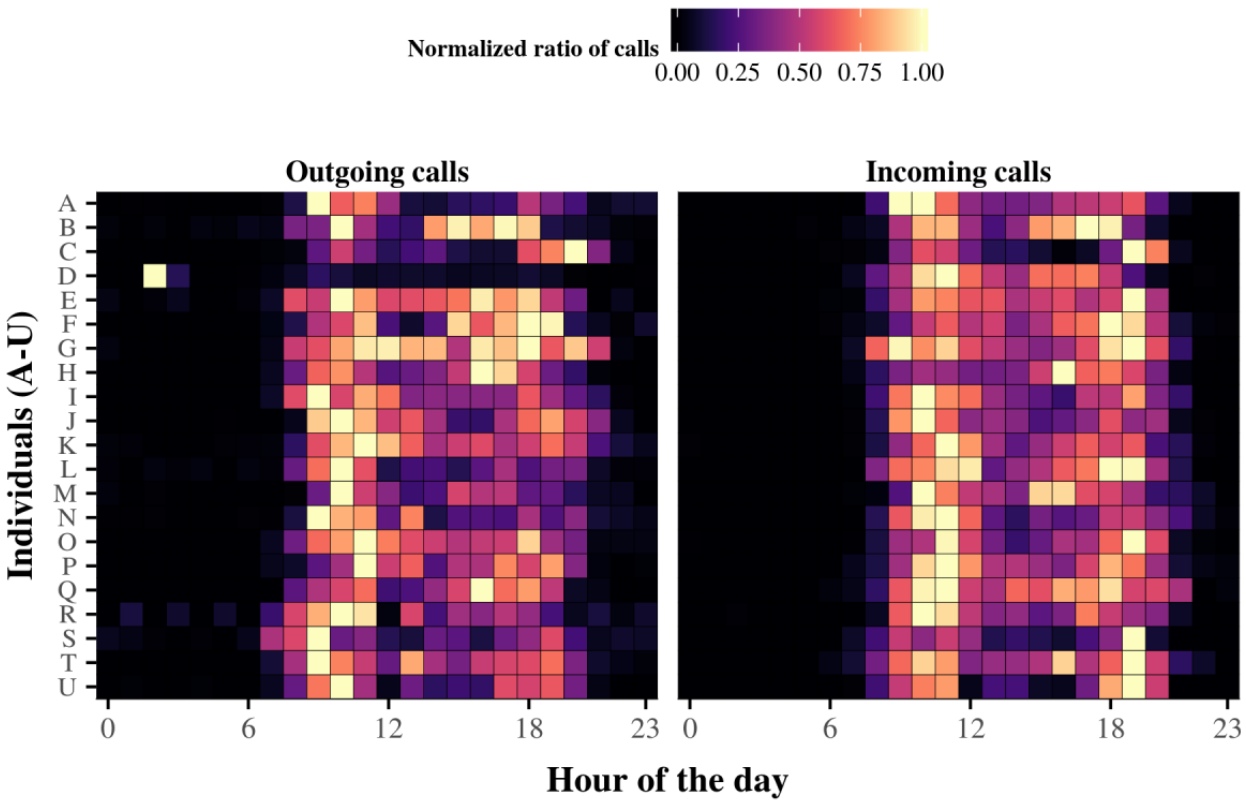
Figure 2. Representative daily rhythms of outgoing and incoming telephone activity for four egos. The solid (dotted) curve shows the outgoing (incoming) call pattern. When the fraction of outgoing telephone calls is greater than the fraction of incoming telephone calls, the difference is shown in green. In the opposite case, the difference is shown in red. The results suggest that both similarities and differences exist in the way the egos place telephone calls and respond to telephone calls during the day. The importance of these observations also appears to vary between egos. In particular, ego D has a prominent nocturnal outgoing telephone call activity that does not appear in the incoming telephone call activity.



To put these illustrative results into context, Figure 3 shows the normalized circadian patterns of the outgoing and incoming telephone calls at the ego level. Here, for each ego and to avoid visual distraction and the influence of extreme values, we normalized the hourly fraction of telephone calls in the range (0, 1), where unity is the maximum fraction and zero is the minimum fraction (see Statistical Tools in the Methods section). The normalized results are displayed on a comparative heat map to illustrate the existence of various daily rhythms in the older adults participating in this study. Interestingly, for some egos, variations appeared between the outgoing and incoming peaks of telephone activity. For instance, individual D switched from a nocturnal peak in outgoing telephone activity to a morning peak in incoming telephone activity. Similarly, the outgoing telephone activity for egos S, T, and U peaked in the morning, whereas the incoming telephone activity peaked in the evening. These differences imply that the hours most propitious for social interaction depend on the social activity or social reactivity of the ego. Taken together, these observations emphasize the interesting characteristics of telephone activity, such as its potential to reflect the daily rhythms of social interactions in

older people and to reveal possible discrepancies between social activity and reactivity.

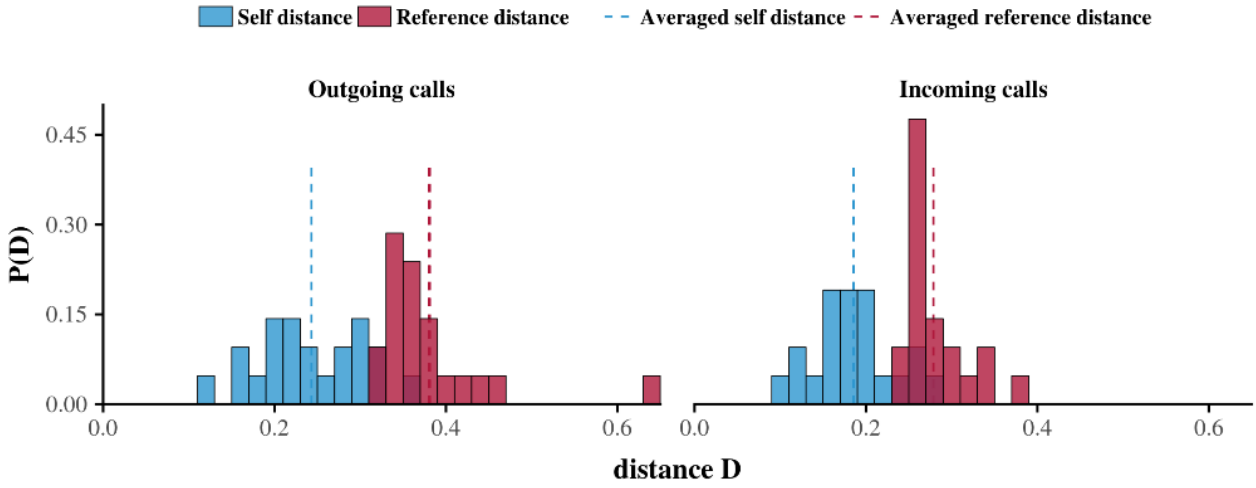
Figure 3. Normalized daily rhythms of outgoing and incoming telephone activity at the individual level. The normalized daily rhythms for each ego are aggregated into a heat map for the outgoing telephone calls (right panel) and for the incoming telephone calls (left panel). The colored squares in each row give the daily rhythm for the telephone calls of one ego with 1-hour resolution (see color scale at the top). Higher fractions of telephone calls correspond to brighter squares. The results show that, in general, most telephone activity occurs during the day rather than at night, except for ego D. Similarities and differences also appear in the way egos place telephone calls versus how they receive telephone calls during the day. Taken together, these observations highlight the potential of telephone activity to reveal the daily rhythms of social activity in older adults.



Evidence of Persistent Daily Rhythms in Both Outgoing and Incoming Telephone Calls in Older Adults

An approach based on the notion of persistence may be applied to analyze outgoing and incoming telephone calls. This approach was first introduced by Saramäki et al to analyze social signatures [44] and then used by Aledavood et al [32-34] to analyze circadian rhythms. At the individual level, the persistence of the daily rhythms calculated above was assessed by comparing the repeatability in time of each ego's daily patterns (denoted D_{self}) with a reference scale (denoted D_{ref}) separately for the outgoing and incoming telephone calls (see Data Analysis in the Methods section). The results revealed the persistence of daily rhythms. In fact, for each ego, the average self-distance D_{self} was less than the reference distance D_{ref} for outgoing telephone calls. Similarly, the average D_{self} was less than D_{ref} for 20 of the 21 egos. Averaging the results for all egos provided $\langle D_{self} \rangle$ = 0.24 (SD 0.06) and $\langle D_{ref} \rangle$ = 0.38 (SD 0.07) for outgoing telephone calls, and $\langle D_{self} \rangle$ = 0.19 (SD 0.03) and $\langle D_{ref} \rangle$ = 0.28 (SD 0.04) for incoming telephone calls (Figure 4).

Figure 4. Average persistence histogram. Red bars represent the average reference distances, whereas blue bars represent the average self-distances for all egos in this study. Blue and red dashed lines represent the average reference distance and the average self-distance of the overall population, respectively. The results show that, on average, the self-distance of egos is less than their reference distance, which is evidence of circadian rhythms in telephone activity among older adults.



Additionally, we confirmed the persistence analysis by using a nonparametric statistical approach. To this end, for each ego, we compared the successive daily patterns of telephone call activity by applying Kolmogorov-Smirnov comparison tests to outgoing and incoming telephone calls separately, as previously performed [32]. The *P* values were strictly greater than .05 for 40 out of 42 cases concerning outgoing telephone calls, and after applying the Holm-Sidak method to correct for multiple comparisons [38], the adjusted *P* values were greater than .05 for all 42 cases. Similarly, for the incoming telephone calls, we obtain *P* values greater than .05 and adjusted *P* values greater than .05 for all 42 cases.

Taken together, these results indicate that, in general, the similarity in the egos' daily rhythms cannot be rejected. Alternatively, a similar persistence analysis may be performed by using a classic Euclidean distance (L^2) instead of the Jensen-Shannon divergence measure, as proposed previously [32]. Executing this alternative led to a similar conclusion for the rhythms of both the outgoing and incoming telephone calls (Table 2).

Table 2. Summary of averaged square root of Jensen-Shannon divergence and L2 metrics for the overall population for both reference distances and self-distances.

Variable	√JSD[a] distance	L^{2b} distance
Self-distance, mean (SD)		
Outgoing	0.24 (0.06)	0.14 (0.04)
Incoming	0.19 (0.03)	0.14 (0.06)
Reference distance, mean (SD)		
Outgoing	0.38 (0.07)	0.23 (0.07)
Incoming	0.28 (0.04)	0.24 (0.11)

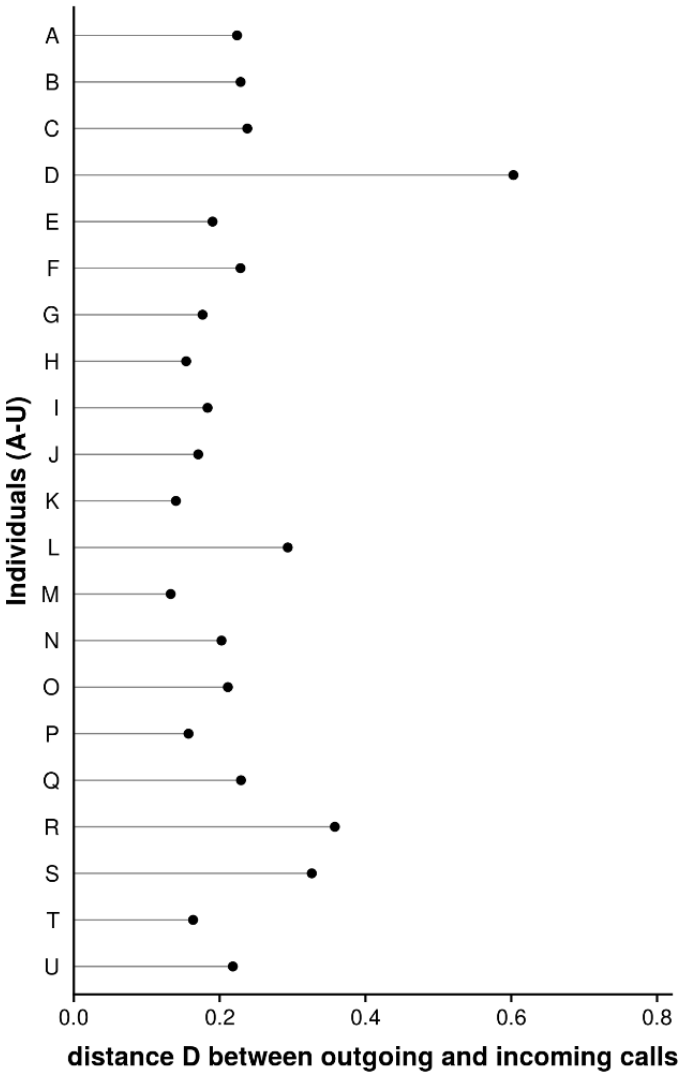
[a]JSD: Jensen-Shannon divergence.

[b]L^2: classic Euclidean distance.

Synchronization of Daily Rhythms for Outgoing and Incoming Telephone Calls at the Individual Level

Having thus far seen evidence that persistent daily rhythms exist in outgoing and incoming telephone calls and that these rhythms may correspond to circadian rhythm framing, we investigated any possible synchronization between these two daily rhythms. First, Figure 5 shows the square root of Jensen-Shannon divergence (see Statistical Tools in the Methods section) averaged over the 12-month study period for each ego to illustrate the synchronization between outgoing and incoming telephone calls at the individual level. The results revealed different dissimilarity values between the distributions of the outgoing and incoming telephone calls, which, except for one ego, tended to remain below a certain threshold.

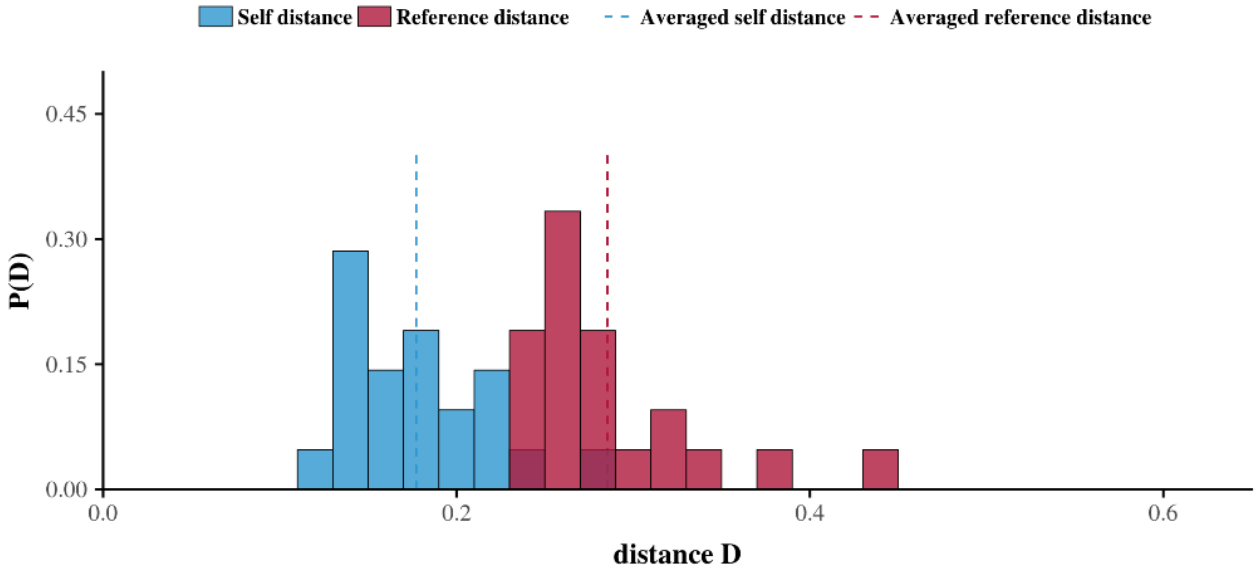
Figure 5. Distance D (ie, square root of Jensen-Shannon divergence) between the distributions of the outgoing and incoming telephone calls at the individual level and over the 12-month study period. The descriptive results show that dissimilarities exist between the daily rhythms for outgoing and incoming telephone calls of the various egos. In particular, although the daily rhythms of outgoing and incoming telephone call activity are highly dissimilar for ego D, the other individuals have dissimilarities that remain under a lower threshold.



We next applied a statistical approach based on the Kolmogorov-Smirnov comparison test under the null hypothesis "H_0: the distributions of outgoing and incoming telephone calls follow the same law." The results emphasized the similarity between these two distributions with a *P* value strictly greater than .05 for all 42 cases. These results can be explained by noting that, despite the variations between the daily rhythms of the outgoing and incoming telephone calls at the individual level, older people tended to distribute their outgoing and incoming communication time in the same way during the day. In short, a notable daily synchronization occurs in the way individuals place and receive telephone calls.

One interesting consequence of this synchronization is that, at the individual level, when combining the outgoing and incoming telephone calls without distinguishing their direction, the resulting mixed daily pattern persists over time. In fact, to evaluate the persistence of these combined rhythms over time, we estimated the daily rhythms of the telephone call activity of each ego by merging together their outgoing and incoming telephone calls and removing the call direction. The results showed that, on average, 20 of the 21 individuals were persistent in their daily rhythms of telephone call activity. In addition, a statistical approach confirmed these results by comparing, for each ego and for each time period, the successive daily patterns of telephone call activity. Applying the Kolmogorov-Smirnov comparison test provided *P* values greater than .05 for all 42 cases (Figure 6). Taken together, these results show that, because of the synchronization, the daily rhythms of outgoing and incoming telephone activity can be merged without distinction. However, these results do not imply that the daily rhythms for outgoing and incoming telephone activity contain exactly the same information about how older people communicate with their social network during the day. Instead, older people have similar rhythms in their telephone activity and reactivity. Further, the alter specificity investigation designed by Aledavood et al [32] should provide useful insights.

Figure 6. Histogram showing the average persistence for mixed outgoing-incoming telephone calls. Red bars represent the averaged reference distances, whereas blue bars represent the averaged self-distances for all the egos in the study. Blue and red dashed lines represent the averaged self-distance and the averaged reference distance of the overall population in the study, respectively. The results suggest that, on average, individuals have a self-distance that is less than their reference distance, which implies that, when mixing outgoing and incoming telephone calls without distinction, the daily rhythms of older adults persist over time.



Alter Specificity Similarities in the Daily Rhythms of Outgoing and Incoming Telephone Calls

The results presented above show that persistent rhythms exist at the individual level for both outgoing and incoming telephone calls and that these rhythms are clearly synchronized with respect to the direction of the telephone call. However, these results do not provide alter specificity information about how older adults interact during the day with their social network. To clarify this situation, we used the method of Aledavood et al [32] to investigate whether alter specificity exists separately for outgoing and incoming telephone calls (see Data Analysis in the Methods section).

Overall, comparing the circadian rhythm of the outgoing telephone calls with that of the incoming telephone calls suggested that similarities exist for alter specificity. For both calling directions, the average relative entropy (see Data Analysis in the Methods section) of the population shown in Figure 7A indicated that egos tend to have low entropy in the evening and night but higher entropy in the morning and afternoon. However, this result did not hold for all egos, as shown in Figure 7B. Additionally, the incoming circadian rhythms at night nuance the relevance of these results because only four egos were active during this time interval. This low number of observations prevented us from concluding that the average entropy for this time period applies to the whole population.

We next investigated whether particular rhythms exist at the individual level in communications involving the favorite telephone contacts of egos for both outgoing and incoming telephone calls. Figure 8 shows the specificity in telephoning to the top two contacts (ie, the two most-contacted alters), as performed by Aledavood et al [32]. The results suggest that similarities exist in how egos place telephone calls and respond to telephone calls with their two favorite alters during the day. Specifically, for outgoing and incoming telephone calls, the corresponding average pattern gives the maximum fraction of calls with the top two contacts in the evening and night, whereas it decreases in the morning and afternoon, as shown in Figure 8A. Additionally, this average fraction of calls also seems inversely proportional to the average relative entropy (Figure 7A). The heat map for the top two alters shown in Figure 8B confirmed that a large fraction of the telephone communications with the top two alters occurred at night. However, as discussed for Figure 7, because only four egos were contacted at night, we cannot extrapolate the resulting average entropy to the whole population.

Figure 7. Relative entropy at the individual level. Panel A shows the averaged relative entropy for both outgoing (black curve) and incoming (dashed curve) telephone calls. Panel B summarizes the relative entropy of all 21 egos in a heat map. Light (dark) colors correspond to low (high) entropy, and gray indicates missing values. Although individuals have their own alter specificity during the daytime, the average relative entropy of the population is similar, except at nighttime when too much data are missing to obtain a significant result.

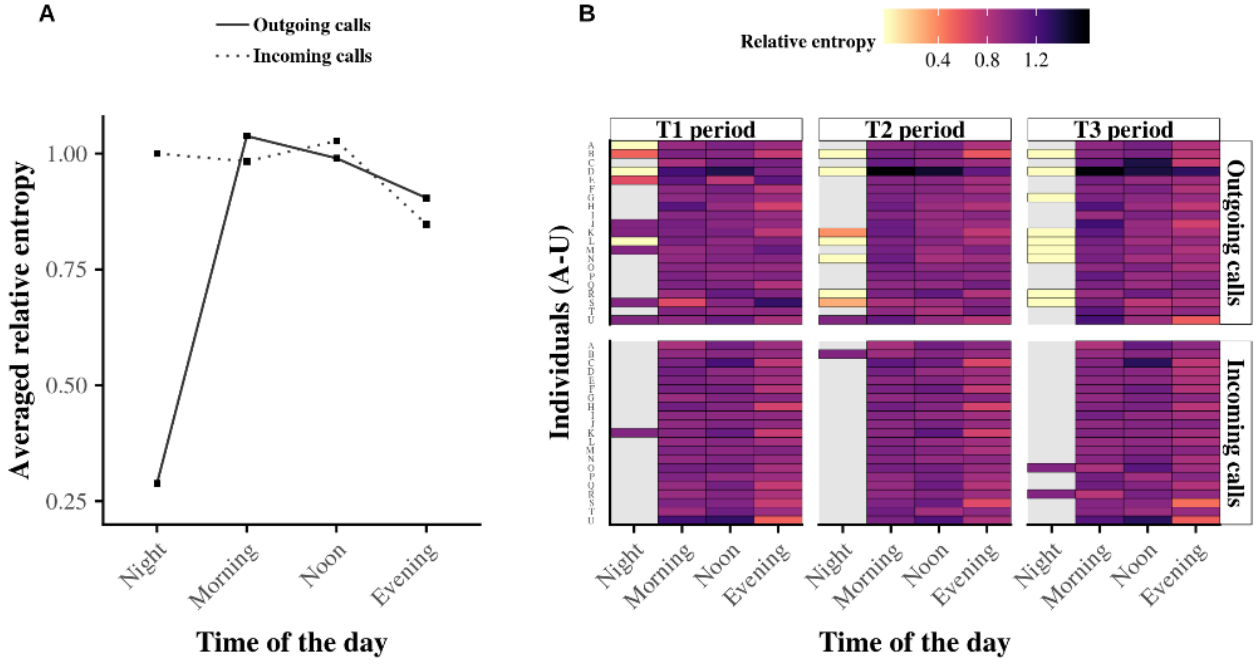
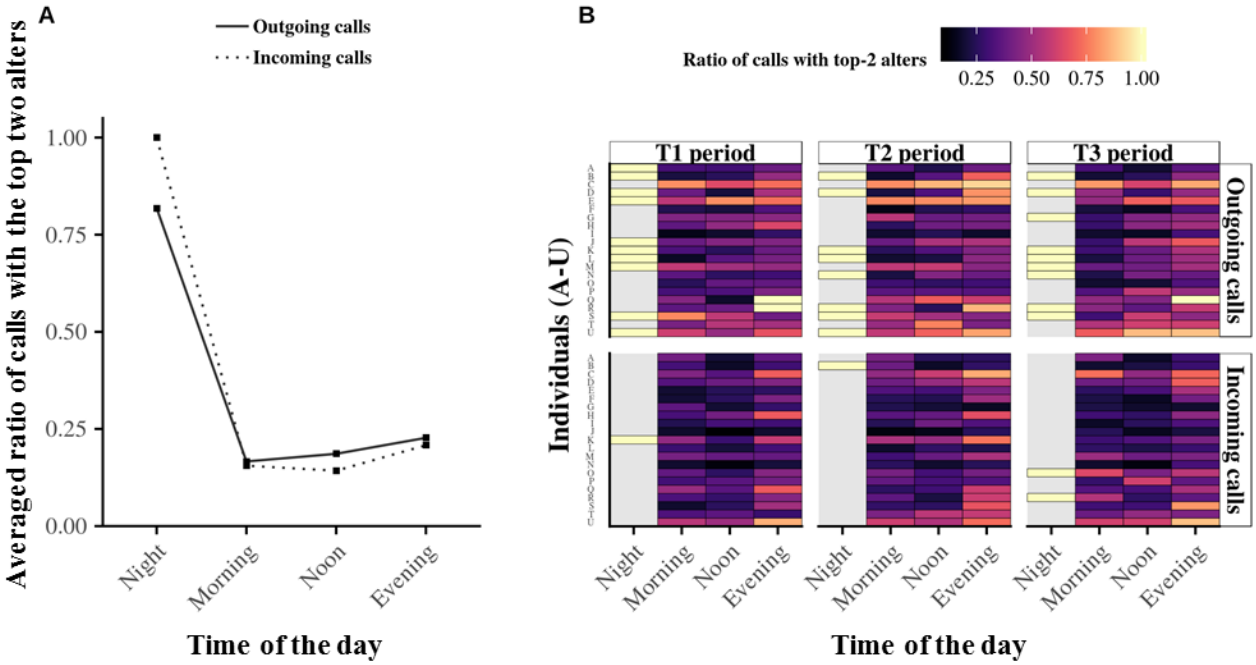
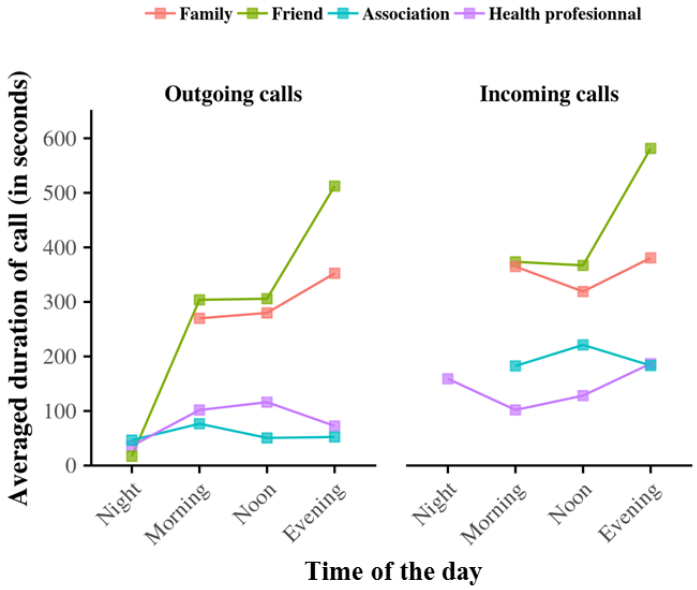


Figure 8. Fraction of telephone calls with the top two alters at the individual level. Panel A shows the average fraction of calls with the top two alters for both outgoing (black curve) and incoming (dotted curve) telephone calls. Panel B summarizes on a heat map the fraction of telephone calls with the top two alters for all 21 egos. A light (dark) color indicates a high (low) fraction, and gray indicates missing values. Although individuals have their own top two alter specificity during the day, the average relative entropy of the population seems similar, except at nighttime when too much data are missing to draw significant conclusions.



Finally, we completed our investigation into alter specificity by considering the duration of telephone calls between egos and their social network during the day. Figure 9 shows the results. The following two results clearly stand out: (1) the duration of telephone calls tended to increase throughout the day until the night, when it dropped rapidly to a minimum and (2) on average, egos communicated longer with their close social network (family and friends) than with others (associations and health professionals). In fact, a threshold in call duration seemed to separate these two social circles.

Figure 9. Average duration of telephone calls between egos and their social network during the day. Data show the average duration of telephone conversations between egos and their family (red), friends (green), health professionals (purple), and associations (blue).



Discussion

Principal Findings

Persistence and synchronization are important mechanisms that affect the daily rhythms of the social activities of older individuals. Unfortunately, they remain underinvestigated in the current literature on call detail record analysis. This work contributes to filling this gap by comparing estimates of circadian rhythms in call detail records of outgoing and incoming telephone calls of an older population. To this end, we focused on the following two issues: (1) the possibility of circadian rhythms existing in the outgoing and incoming telephone calling activities of seniors, and (2) the possible existence of synchronization between these two rhythms. This study used a data set of 12 successive months that combined call detail records and questionnaire data from 21 volunteers, all of whom were over 70 years of age. In particular, we extracted the following four specific telecommunication parameters from the call detail records: (1) recipient of the call (ie, alter), (2) time at which the call began, (3) duration of the call, and (4) direction of the call (outgoing or incoming).

The results indicated that differences exist between the daily rhythms of the salient features of outgoing and incoming telephone activities (eg, hourly peaks of calling) by older individuals (Figures 1-4) and that both of these rhythms are in accordance with circadian rhythm framing. Despite these differences, the daily rhythms of outgoing and incoming telephone calls generally follow the same distribution. In other words, the rhythms are synchronized. Interestingly, one consequence of this synchronization is that the daily rhythm persists upon combining outgoing and incoming telephone calls (Figure 6). Finally, based on a descriptive approach, we suggest that, behind this temporal synchronization, individuals also present similarities in the way that they place calls to and receive calls from specific alters during specific periods of the day (Figures 7-9).

Comparing the present findings on older adults with the existing literature on young individuals [32,33] reveals similarities in the way in which old and young individuals make telephone calls during the day. Both older adults and younger individuals exhibit consistent daily rhythms in their telephone activity in accordance with circadian rhythm framing. Furthermore, for both younger and older individuals, telephone calls made in the evening and night were more focused on specific well-known alters than those made during the rest of the day. However, particular differences also existed between these two populations. In particular, at the aggregate level, older adults exhibited general circadian rhythms of telephone activity with a bimodal distribution. Their telephone activity peaked in the morning and again in the evening, separated by a period of low activity in the afternoon. The circadian rhythm of the telephone activity of younger individuals, however, tends to follow a Gaussian distribution centered on the afternoon period [32,33].

This important difference between younger and older individuals may be explained by invoking the following four hypotheses:

(1) With increasing age, people tend to nap more in the afternoon, which could explain the period of low telephone activity in the older population studied herein [48].

(2) The work schedule for younger adults, as noted previously [32,33], may module their social activity and thus their telephone activity. Thus, specific periods of the day, such as after school, could be more propitious for telephone activity than other periods, such as the morning.

(3) In the older population, the chronotype shift that usually appears with age tends to make people more active in the morning, which could explain the first peak of telephone activity in the morning [17].

(4) Finally, cultural differences between the two samples, such as the country of origin of the participants (France in this study and the United Kingdom in previous studies [32,33]) may also impact the general circadian rhythms observed in the two studies, notably regarding cultural zeitgebers such as mealtimes and nap times.

Thus, taken together, the present results show that (1) older individuals may present persistent daily rhythms in telephone activity for both outgoing and incoming telephone calls in accordance with circadian rhythm framing and (2) the daily rhythm of their outgoing and incoming calls may be synchronized. These circadian rhythms contain both similarities and differences when compared with those of other populations, such as younger adults. In the field of health monitoring, these results suggest that the telephone may serve as a digital sensor of the temporal social activity of older individuals. In clinical practice, the telephone activity of older patients could be monitored to supervise their health in real-time based on their social activity and reactivity. Such supervision could consist of detecting particular health issues in progress, such as insomnia and other sleep disturbances that often occur with age [17], which could be signaled by nocturnal telephone activity. Additionally, such supervision could help to prevent risky social situations, such as isolation and depression, by detecting a decrease in the number and duration of outgoing and incoming telephone calls with individuals in the close social network of the patient. Both of these detections (ie, health issues and social risks) could be presented to health professionals in the form of triggers to help them better understand and anticipate health complications or worsening conditions of older patients. Notably, information stemming from monitoring telephone activity has the advantage of being collected in a passive and unobtrusive way. Such monitoring could be done in conjunction with traditional clinical questionnaires about the rhythms of social activity that are often used in clinical practice to monitor the daily behavior of patients [18]. Thus, this digital approach could potentially improve traditional health care systems by providing objective estimates of the social activity of patients to augment the subjective answers to questionnaires collected at more dispersed time intervals. However, before considering such an approach, the limitations described in the next section should be considered.

Limitations

Because we provided a descriptive analysis based on a relatively rich but small sample of 21 older individuals, no straightforward generalization of the results should be made to the overall elderly population. In particular, given the potential sparsity of data and the small number of telephone calls made each day by the subjects in this study (Table 1), no foregone conclusions should be made regarding observations made under different conditions or on different data sets. To resolve this, future work should consider larger data sets and, more broadly, analyze different populations, such as healthy, disabled, or chronically ill individuals, or even vary the time period of observation. An evaluation of the reproducibility of the proposed approach would help to determine the robustness of the present results, which would in turn determine the feasibility of analyzing call detail records for health-monitoring purposes.

In addition, given the small sample size, the present quantitative data should be confronted with qualitative data. For health monitoring, such an approach could provide an interesting perspective to contextualize the present findings.

Finally, a technical limitation is present in the size of the time interval used in our data analysis for investigating the persistence and synchronization of circadian rhythms over time. Although no clear consensus exists in the literature about the size of the time interval when analyzing temporal telephone data, previous studies [41] have reported that aggregating data over long time intervals (eg, several months) limits the effect of short-term variations in call patterns, thereby enhancing the robustness of the persistence analysis. Consequently, by splitting our 12-month data set into three equal 4-month intervals as performed in previous work [32], the results for persistence and synchronization are valid only for a 4-month interval. For other intervals, further studies are required. Such investigations are welcome because call detail record analyses generally suffer from a lack of investigation into the reproducibility of the existing methodology to more general conditions, as mentioned previously [44].

Conclusion

In this work, we estimated the circadian rhythms of the telephone activity of older adults. The results showed (1) temporal persistence and (2) synchronization between how older adults place and receive telephone calls over the day. At this point, more investigations are needed to determine the extent to which such rhythms may be harnessed for health-monitoring purposes. In particular, the results of this study showed that, despite high persistence and synchronization, differences exist between the daily rhythms of the outgoing and incoming telephone activities of older adults. These differences may relate, for example, to the daily peak hour in telephone activity.

Perspectives

Determining the similarities and dissimilarities between the daily rhythms of outgoing and incoming telephone activities for health-monitoring purposes is of interest to health care professionals. At this point, however, the methodology used here, which is largely inspired by a previous report [32], does not permit us to determine the significance of the results. Along these lines, extracting relevant and important information from telephone activity for health monitoring is part of our immediate plan. We believe that such an approach could provide an opportunity for assessing the relevance of digital circadian rhythms to characterize the behavior of older patients and, more broadly, to investigate the role of new technologies in improving health care monitoring. Future studies should investigate, in particular, older adults with onset of neurodegenerative diseases, such as Alzheimer disease, for whom the persistence evinced here could be replaced by a transition to a pathologic perseveration that could be detected and characterized by repeated outgoing calls with a given circadian rhythm.

Conflicts of Interest

None declared.

References

1. Honma S. The mammalian circadian system: a hierarchical multi-oscillator structure for generating circadian rhythm. J Physiol Sci 2018 May;68(3):207-219. [doi: 10.1007/s12576-018-0597-5] [Medline: 29460036]
2. Czeisler C, Duffy J, Shanahan T, Brown E, Mitchell J, Rimmer D, et al. Stability, precision, and near-24-hour period of the human circadian pacemaker. Science 1999 Jun 25;284(5423):2177-2181 [FREE Full text] [doi: 10.1126/science.284.5423.2177] [Medline: 10381883]
3. Elmore SK, Betrus PA, Burr R. Light, social zeitgebers, and the sleep-wake cycle in the entrainment of human circadian rhythms. Res Nurs Health 1994 Dec;17(6):471-478. [doi: 10.1002/nur.4770170610] [Medline: 7972925]
4. Ehlers CL, Frank E, Kupfer DJ. Social zeitgebers and biological rhythms. A unified approach to understanding the etiology of depression. Arch Gen Psychiatry 1988 Oct 01;45(10):948-952. [doi: 10.1001/archpsyc.1988.01800340076012] [Medline: 3048226]
5. Monk TH, Flaherty JF, Frank E, Hoskinson K, Kupfer DJ. The Social Rhythm Metric. An instrument to quantify the daily rhythms of life. J Nerv Ment Dis 1990 Feb;178(2):120-126. [doi: 10.1097/00005053-199002000-00007] [Medline: 2299336]
6. Goldbeter A. La Vie oscillatoire: Au coeur des rythmes du vivant. Paris, France: Jacob; 2018.
7. Grandin LD, Alloy LB, Abramson LY. The social zeitgeber theory, circadian rhythms, and mood disorders: review and evaluation. Clin Psychol Rev 2006 Oct;26(6):679-694. [doi: 10.1016/j.cpr.2006.07.001] [Medline: 16904251]
8. Smagula SF, DuPont CM, Miller MA, Krafty RT, Hasler BP, Franzen PL, et al. Rest-activity rhythms characteristics and seasonal changes in seasonal affective disorder. Chronobiol Int 2018 Oct 19;35(11):1553-1559 [FREE Full text] [doi: 10.1080/07420528.2018.1496094] [Medline: 30024782]
9. Wittmann M, Dinich J, Merrow M, Roenneberg T. Social jetlag: misalignment of biological and social time. Chronobiol Int 2006 Jul;23(1-2):497-509. [doi: 10.1080/07420520500545979] [Medline: 16687322]
10. McMahon DM, Burch JB, Wirth MD, Youngstedt SD, Hardin JW, Hurley TG, et al. Persistence of social jetlag and sleep disruption in healthy young adults. Chronobiol Int 2018 Mar 12;35(3):312-328 [FREE Full text] [doi: 10.1080/07420528.2017.1405014] [Medline: 29231745]
11. Levandovski R, Dantas G, Fernandes LC, Caumo W, Torres I, Roenneberg T, et al. Depression Scores Associate With Chronotype and Social Jetlag in a Rural Population. Chronobiology International 2011 Sep 06;28(9):771-778. [doi: 10.3109/07420528.2011.602445] [Medline: 21895489]
12. Malone S, Mendoza M, Patterson F. Social jetlag, circadian disruption, and cardiometabolic disease risk. In: Sleep and Health. Cambridge, Massachusetts, United States: Academic Press; 2019:227-240.
13. Chen C, Logan RW, Ma T, Lewis DA, Tseng GC, Sibille E, et al. Effects of aging on circadian patterns of gene expression in the human prefrontal cortex. Proc Natl Acad Sci U S A 2016 Jan 05;113(1):206-211 [FREE Full text] [doi: 10.1073/pnas.1508249112] [Medline: 26699485]
14. Duffy JF, Zitting K, Chinoy ED. Aging and Circadian Rhythms. Sleep Med Clin 2015 Dec;10(4):423-434 [FREE Full text] [doi: 10.1016/j.jsmc.2015.08.002] [Medline: 26568120]
15. Cornelissen G, Otsuka K. Chronobiology of Aging: A Mini-Review. Gerontology 2017 Oct 22;63(2):118-128 [FREE Full text] [doi: 10.1159/000450945] [Medline: 27771728]
16. Steponenaite A, Biello SM, Lall GS. Aging clocks: disrupted circadian rhythms. Aging (Albany NY) 2018 Nov 13;10(11):3065-3066 [FREE Full text] [doi: 10.18632/aging.101642] [Medline: 30425185]
17. Monk TH, Buysse DJ. Chronotype, bed timing and total sleep time in seniors. Chronobiol Int 2014 Jun 11;31(5):655-659 [FREE Full text] [doi: 10.3109/07420528.2014.885981] [Medline: 24517139]
18. Monk T, Reynolds C, Machen M, Kupfer D. Daily social rhythms in the elderly and their relation to objectively recorded sleep. Sleep 1992 Aug;15(4):322-329. [doi: 10.1093/sleep/15.4.322] [Medline: 1519007]
19. Lowenthal MF. Social Isolation and Mental Illness in Old Age. American Sociological Review 1964 Feb;29(1):54. [doi: 10.2307/2094641]
20. Domènech-Abella J, Lara E, Rubio-Valera M, Olaya B, Moneta MV, Rico-Uribe LA, et al. Loneliness and depression in the elderly: the role of social network. Soc Psychiatry Psychiatr Epidemiol 2017 Apr 2;52(4):381-390. [doi: 10.1007/s00127-017-1339-3] [Medline: 28154893]
21. Aselage M, Amella E. An evolutionary analysis of mealtime difficulties in older adults with dementia. J Clin Nurs 2010 Jan;19(1-2):33-41. [doi: 10.1111/j.1365-2702.2009.02969.x] [Medline: 20500242]
22. Wehrens SM, Christou S, Isherwood C, Middleton B, Gibbs MA, Archer SN, et al. Meal Timing Regulates the Human Circadian System. Curr Biol 2017 Jun 19;27(12):1768-1775.e3 [FREE Full text] [doi: 10.1016/j.cub.2017.04.059] [Medline: 28578930]
23. Musiek ES, Xiong DD, Holtzman DM. Sleep, circadian rhythms, and the pathogenesis of Alzheimer disease. Exp Mol Med 2015 Mar 13;47(3):e148-e148 [FREE Full text] [doi: 10.1038/emm.2014.121] [Medline: 25766617]

24. Musiek ES, Bhimasani M, Zangrilli MA, Morris JC, Holtzman DM, Ju YS. Circadian Rest-Activity Pattern Changes in Aging and Preclinical Alzheimer Disease. JAMA Neurol 2018 May 01;75(5):582-590 [FREE Full text] [doi: 10.1001/jamaneurol.2017.4719] [Medline: 29379963]
25. Duncan MJ. Interacting influences of aging and Alzheimer's disease on circadian rhythms. Eur J Neurosci 2020 Jan 21;51(1):310-325. [doi: 10.1111/ejn.14358] [Medline: 30689226]
26. Wu JQ, Li P, Stavitsky Gilbert K, Hu K, Cronin-Golomb A. Circadian Rest-Activity Rhythms Predict Cognitive Function in Early Parkinson's Disease Independently of Sleep. Mov Disord Clin Pract 2018 Nov 08;5(6):614-619 [FREE Full text] [doi: 10.1002/mdc3.12692] [Medline: 30637282]
27. Seppälä J, De Vita I, Jämsä T, Miettunen J, Isohanni M, Rubinstein K, M-RESIST Group, et al. Mobile Phone and Wearable Sensor-Based mHealth Approaches for Psychiatric Disorders and Symptoms: Systematic Review. JMIR Ment Health 2019 Feb 20;6(2):e9819 [FREE Full text] [doi: 10.2196/mental.9819] [Medline: 30785404]
28. Laport-López F, Serrano E, Bajo J, Campbell AT. A review of mobile sensing systems, applications, and opportunities. Knowl Inf Syst 2019 Mar 8;62(1):145-174. [doi: 10.1007/s10115-019-01346-1]
29. Jones KH, Daniels H, Heys S, Ford DV. Challenges and Potential Opportunities of Mobile Phone Call Detail Records in Health Research: Review. JMIR Mhealth Uhealth 2018 Jul 19;6(7):e161 [FREE Full text] [doi: 10.2196/mhealth.9974] [Medline: 30026176]
30. Bhugra D, Tasman A, Pathare S, Priebe S, Smith S, Torous J, et al. The WPA- Lancet Psychiatry Commission on the Future of Psychiatry. The Lancet Psychiatry 2017 Oct;4(10):775-818. [doi: 10.1016/S2215-0366(17)30333-4] [Medline: 28946952]
31. Russ TC, Woelbert E, Davis KA, Hafferty JD, Ibrahim Z, Inkster B, MQ Data Science group. How data science can advance mental health research. Nat Hum Behav 2019 Jan 10;3(1):24-32. [doi: 10.1038/s41562-018-0470-9] [Medline: 30932051]
32. Aledavood T, López E, Roberts SG, Reed-Tsochas F, Moro E, Dunbar RI, et al. Daily Rhythms in Mobile Telephone Communication. PLoS One 2015 Sep 21;10(9):e0138098 [FREE Full text] [doi: 10.1371/journal.pone.0138098] [Medline: 26390215]
33. Aledavood T, López E, Roberts S, Reed-Tsochas F, Moro E, Dunbar R. Channel-Specific Daily Patterns in Mobile Phone Communication. In: Battiston S, De Pellegrini F, Caldarelli G, Merelli E, editors. Proceedings of ECCS 2014. Springer Proceedings in Complexity. Cham: Springer International Publishing; 2016.
34. Aledavood T, Lehmann S, Saramäki J. Digital daily cycles of individuals. Front. Phys 2015 Oct 07;3. [doi: 10.3389/fphy.2015.00073]
35. Aledavood T, Lehmann S, Saramäki J. Social network differences of chronotypes identified from mobile phone data. EPJ Data Sci 2018 Oct 24;7(1). [doi: 10.1140/epjds/s13688-018-0174-4]
36. Agrawal R, Pati AK. Multi-frequency rhythms in activity of mixed human population in social networking sites – a preliminary study. Biological Rhythm Research 2017 May 29;49(1):1-11. [doi: 10.1080/09291016.2017.1333566]
37. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study. J Med Internet Res 2015 Jul 15;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
38. Murnane E, Abdullah S, Matthews M, Choudhury T, Gay G. Social (media) jet lag: how usage of social technology can modulate and reflect circadian rhythms. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. Osaka, Japan: ACM Press; 2015:843-854.
39. Cho C, Lee T, Kim M, In HP, Kim L, Lee H. Mood Prediction of Patients With Mood Disorders by Machine Learning Using Passive Digital Phenotypes Based on the Circadian Rhythm: Prospective Observational Cohort Study. J Med Internet Res 2019 Apr 17;21(4):e11029. [doi: 10.2196/11029]
40. Lin Y, Wong B, Pan Y, Chiu Y, Lee Y. Validation of the Mobile App-Recorded Circadian Rhythm by a Digital Footprint. JMIR Mhealth Uhealth 2019 May 16;7(5):e13421 [FREE Full text] [doi: 10.2196/13421] [Medline: 31099340]
41. Aubourg T, Demongeot J, Renard F, Provost H, Vuillerme N. Association between social asymmetry and depression in older adults: A phone Call Detail Records analysis. Sci Rep 2019 Sep 18;9(1):13524 [FREE Full text] [doi: 10.1038/s41598-019-49723-8] [Medline: 31534178]
42. Aubourg T, Demongeot J, Provost H, Vuillerme N. Circadian Rhythms in the Telephone Calls of Older Adults: Observational Descriptive Study. JMIR Mhealth Uhealth 2020 Feb 25;8(2):e12452 [FREE Full text] [doi: 10.2196/12452] [Medline: 32130156]
43. Aubourg T, Demongeot J, Renard F, Provost H, Vuillerme N. How to Measure Circadian Rhythms of Activity and Their Disruptions in Humans Using Passive and Unobtrusive Capture of Phone Call Activity. Stud Health Technol Inform 2019 Aug 21;264:1631-1632. [doi: 10.3233/SHTI190569] [Medline: 31438265]
44. Saramäki J, Leicht EA, López E, Roberts SG, Reed-Tsochas F, Dunbar RI. Persistence of social signatures in human communication. Proc Natl Acad Sci U S A 2014 Jan 21;111(3):942-947 [FREE Full text] [doi: 10.1073/pnas.1308540110] [Medline: 24395777]
45. Blondel VD, Decuyper A, Krings G. A survey of results on mobile phone datasets analysis. EPJ Data Sci 2015 Aug 5;4(1). [doi: 10.1140/epjds/s13688-015-0046-0]
46. Massey FJ. The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association 1951 Mar;46(253):68-78. [doi: 10.1080/01621459.1951.10500769]

47. Dodge Y. The Oxford dictionary of statistical terms. Oxford: Oxford University Press; 2006.
48. Milner C, Cote K. Benefits of napping in healthy adults: impact of nap length, time of day, age, and experience with napping. J Sleep Res 2009 Jun;18(2):272-281 [FREE Full text] [doi: 10.1111/j.1365-2869.2008.00718.x] [Medline: 19645971]

Edited by G Eysenbach; submitted 29.01.19; peer-reviewed by T Aledavood, E Murnane; comments to author 01.05.19; revised version received 26.06.19; accepted 28.01.20; published 26.11.20.

Please cite as:
Aubourg T, Demongeot J, Provost H, Vuillerme N
Exploitation of Outgoing and Incoming Telephone Calls in the Context of Circadian Rhythms of Social Activity Among Elderly People: Observational Descriptive Study
JMIR Mhealth Uhealth 2020;8(11):e13535
URL: http://mhealth.jmir.org/2020/11/e13535/
doi:10.2196/13535
PMID:33242018

Original Paper

Assessing User Retention of a Mobile App: Survival Analysis

Yu-Hsuan Lin[1,2,3,4], MD, PhD; Si-Yu Chen[1], MS; Pei-Hsuan Lin[5], BSc; An-Shun Tai[5], PhD; Yuan-Chien Pan[1], MSc; Chang-En Hsieh[1], BSc; Sheng-Hsuan Lin[5], MD, ScD

[1]Institute of Population Health Sciences, National Health Research Institutes, Miaoli, Taiwan
[2]Department of Psychiatry, National Taiwan University Hospital, Taipei, Taiwan
[3]Department of Psychiatry, College of Medicine, National Taiwan University, Taipei, Taiwan
[4]Institute of Health Behaviors and Community Sciences, College of Public Health, National Taiwan University, Taipei, Taiwan
[5]Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan

Corresponding Author:
Sheng-Hsuan Lin, MD, ScD
Institute of Statistics
National Chiao Tung University
1001 University Road
Hsinchu, 300
Taiwan
Phone: 886 (3) 5712121 ext 56822
Email: shenglin@nctu.edu.tw

Abstract

Background: A mobile app generates passive data, such as GPS data traces, without any direct involvement from the user. These passive data have transformed the manner of traditional assessments that require active participation from the user. Passive data collection is one of the most important core techniques for mobile health development because it may promote user retention, which is a unique characteristic of a software medical device.

Objective: The primary aim of this study was to quantify user retention for the "Staff Hours" app using survival analysis. The secondary aim was to compare user retention between passive data and active data, as well as factors associated with the survival rates of user retention.

Methods: We developed an app called "Staff Hours" to automatically calculate users' work hours through GPS data (passive data). "Staff Hours" not only continuously collects these passive data but also sends an 11-item mental health survey to users monthly (active data). We applied survival analysis to compare user retention in the collection of passive and active data among 342 office workers from the "Staff Hours" database. We also compared user retention on Android and iOS platforms and examined the moderators of user retention.

Results: A total of 342 volunteers (224 men; mean age 33.8 years, SD 7.0 years) were included in this study. Passive data had higher user retention than active data (P=.011). In addition, user retention for passive data collected via Android devices was higher than that for iOS devices (P=.015). Trainee physicians had higher user retention for the collection of active data than trainees from other occupations, whereas no significant differences between these two groups were observed for the collection of passive data (P=.700).

Conclusions: Our findings demonstrated that passive data collected via Android devices had the best user retention for this app that records GPS-based work hours.

(JMIR Mhealth Uhealth 2020;8(11):e16309) doi:10.2196/16309

KEYWORDS

smartphone; passive data, user retention; mobile application; app; survival analysis; work hours

Introduction

In the past decade, smartphones have become nearly ubiquitous. Over 3 billion smartphones have internet subscriptions, and each device has the information processing capacity of supercomputers of the 1990s [1]. Even in areas without easy access to clean water, ownership of a smartphone and rapid access to information have become symbols of modernity [2]. Although most comparable sources of big data are scarce in the

world's poorest nations, mobile phones are a notable exception [3]. Individuals' history of mobile phone use can be used to infer their socioeconomic status [4]. Mobile apps could also help to fill the gap of health inequality, such as that for workers with extremely long work hours.

User retention is defined as the number of initial users who are still active in a given time frame. User retention is usually calculated as the number of those who are still active divided by the total number of users registered [5]. User retention enables designers to understand how often the users return to use the available service of a mobile app. Tracking user retention is both important and difficult, because the majority of mobile app users only use apps over a short timeframe [6] and report that the most important and acceptable components of a mobile app are ease of use and time, with the average time of use being 9.3 seconds. Gathering enough data to make reliable inferences of the retention rate over long timespans is also a challenge. Recently, user retention was adopted as a crucial index to evaluate the effect and utility of mobile apps designed for self-management. Previous studies have shown that users benefit from long-term engagement with an app [7]. However, the lack of studies with large samples using a mobile app may signal a need for additional studies on the potential use of a mobile app to assist individuals in changing self-management behaviors, such as health behaviors [5].

There were over 90 million documented mental health app installations by the end of 2018 [8]. Although mobile apps have been considered a feasible and acceptable means of administering health intervention, most literature regarding health apps has focused on preventing and managing chronic disease, self-monitoring of health behaviors, or content analyses of health and fitness apps [5]. However, studies investigating the utility of mobile apps as a health intervention were mostly executed in empirical study settings. Understanding patterns of real-world usage of health apps is key to maximizing their potential to increase self-management of care by the public [8]. Although the number of app installs and daily active minutes of use may seem high, only a small portion of users actually use apps for a long period of time [8]. A usage analysis of user engagement of unguided mental health apps found the general user retention is poor, with a median 15-day retention of 3.9% and 30-day retention of 3.3% [8]. Therefore, how to enhance long-term engagement of real-word usage of health apps may be crucial for both designers and users.

Passive data are defined as data that are generated without any direct involvement from the subject, such as GPS traces and phone call logs. By contrast, active data are defined as data that require active participation from the subject, such as surveys and audio samples [9]. Passive data generated by mobile apps have transformed the manner of traditional assessments based on active data. In 2003, the Accreditation Council for Graduate Medical Education implemented work hour limits for all physicians in training in the United States. However, surveying medical interns' compliance with these work hour limits using a traditional assessment took 2 years, and the national survey was published in 2006 [10]. In addition, these self-reports could not reflect fluctuations in work hours in a timely manner, especially for medical staff with frequent on-call duties. Nowadays, smartphones offer us objective and ecological sources of measurement that continuously and passively collect data. These reliable, quantitative data could facilitate real-time policy evaluation and target resources to those with the greatest requirement, even in remote and inaccessible regions. In addition, policy regarding resident physician work hours has shifted frequently in recent years [11,12]; assessing user retention of an app like "Staff Hours" is useful to track the implementation of a work hours policy, as well as the resident physicians' compliance with work hour limits.

We developed an iOS and Android smartphone app called "Staff Hours" that automatically calculates users' work hours through GPS data. This passive data collection by "Staff Hours" is similar to that of our previous apps, "Know Addiction" and "Rhythm," which collect and calculate smartphone screen time and sleep time, respectively [13-18]. However, these two apps were experimentally used for months, and user retention of an app collecting passive data in a natural setting is still unknown. The inherently dynamic nature of apps adds to the challenge of developing reliable metrics of app users' retention. In addition, mobile app development is largely consumer-led and commercial-driven, and the evaluation of user retention is often app-centered and not user-centered. For example, a study tracking the longitudinal availability of mental health apps reported that these apps have a half-life; after a certain amount of time, an app may no longer be available for public use [19]. These app-centered metrics may have a low correlation with the apps' clinical utility or usability [20].

Despite the ability of passive data collection by smartphone apps, little is known about user retention in the real world. Of particular interest are potential factors associated with the user retention of mobile apps. We hypothesized that user retention for passive data collection is higher than that for active data collection. We further hypothesized that power-saving operating systems and the target audience increase user retention for passive data and active data, respectively. The primary aim of this study was to quantify the user retention for the "Staff Hours" app using survival analysis. The secondary aim was to compare user retention between passive data and active data, as well as factors associated with the survival rates of user retention.

Methods

Participants

We collected data from 421 office workers from August 2018 to March 2019 from the "Staff Hours" database, which is owned by the National Health Research Institutes. This newly developed "Staff Hours" app automatically estimates users' work hours daily through a GPS record using an algorithm. All participants were volunteers who were interested in their work hours and had successfully installed this app. We defined any data uploaded to our server during the first 28 days of their registration as successful installation of this app. We excluded 79 of the 421 app users who did not provide demographic data such as gender and age. Otherwise, app users had to provide their demographic data and occupation. A total of 342 participants (224 men; mean age 33.76 years, SD 7.01 years, range 20-59 years) were included in this study. Most of the

participants (286/342, 83.6%) were medical staff, and most of the medical staff (128/286, 44.8%) were trainee physicians (resident physicians). The "Staff Hours" app is only available in Taiwan, and consent was obtained from all users to allow their data to be collected electronically before installation. Different versions of this app were available on the Android or iOS platform. The study was approved by the Institutional Review Board of the National Health Research Institutes. All clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki.

Measures

Designing the App — Staff Hours

The "Staff Hours" app is designed to automatically record GPS data in the background without interrupting the smartphone operating system. In the beginning of installation, users have to provide at least one workplace location; the app can track up to 5 locations simultaneously. The location is transformed to longitude and latitude with the Google map format. The process of recording the GPS-defined work hours is illustrated in Figure 1. The GPS detection range is a 1-km radius around the workplace, and the recording of work hours starts when the workplace location is within range for a consecutive 30 minutes. Similarly, the recording of work hours ends when there is a consecutive 30-minute period without the workplace within 1 km of the device. To save battery power, the sampling rate of GPS data is fixed at 10 minutes. The user interface of the "Staff Hours" app is shown in Figure 2. In addition to GPS-defined work hours, users can set up their work schedules in this app. The user-inputted "scheduled work hours" include "regular work hours" and "on-call duty hours." Aside from total work hours, the app also provides the output "overtime work hours" using the formula of "total work hours" minus "scheduled work hours."

Figure 1. Recording process of GPS-defined work hours. In this example, the user walks into and leaves the 1-km radius centered on the workplace at 8:50 AM and 6:20 PM, respectively. The app then automatically generates GPS-defined work hours in real time (8:50-6:20).

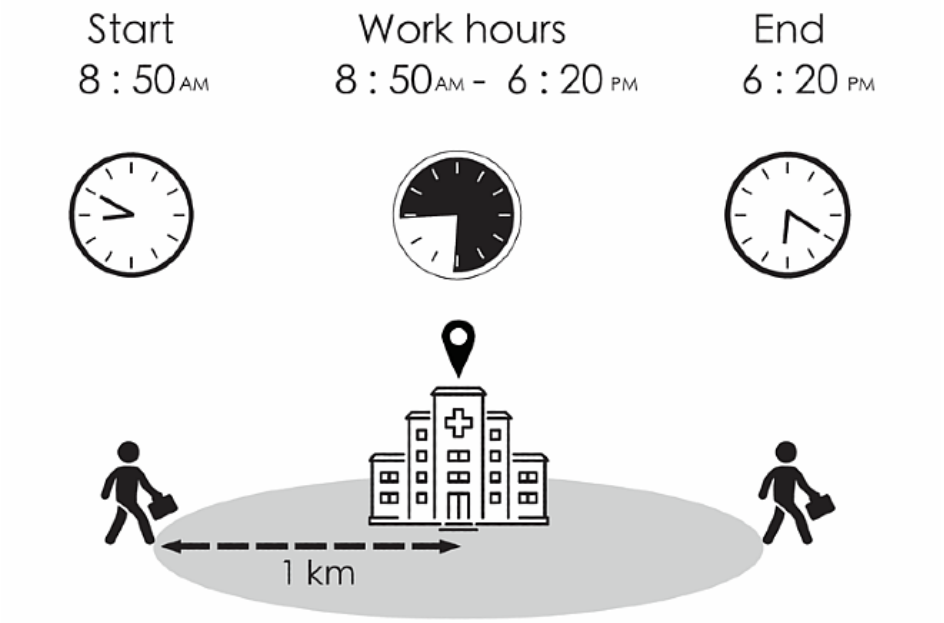
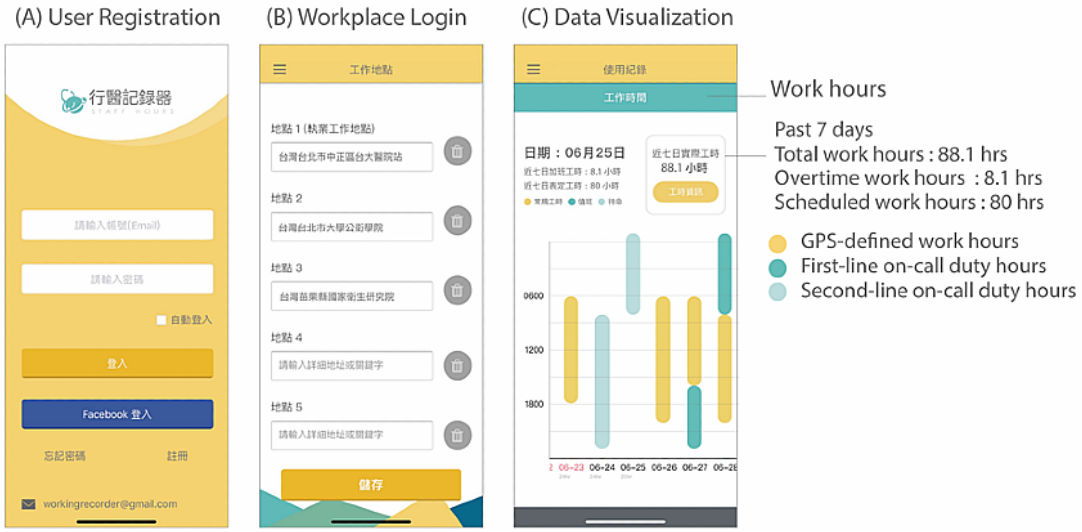


Figure 2. Screenshots of the "Staff Hours" app: (A) home screen and registration; (B) workplace login; (C) data visualization.



Collecting Passive Data — GPS-Defined Work Hours

Collecting and uploading data to the dedicated lab server occur automatically in the app background, and no interference with the daily routine of the smartphone use is observed. “Staff Hours” collects these GPS-defined work hours in the Android operating system, but the app stops during background refresh in the iOS. These GPS-defined work hours data are called “passive data” throughout this article since the data collection does not require any active participation from the smartphone user.

Collecting Active Data — Mental Health Survey

Every morning, there is a notification to remind the app users to examine their work hour records. On the first day of every month, this app pushes a notification for an 11-item mental health survey. This questionnaire includes the Patient Health Questionnaire (PHQ-9) for depressive symptoms [21], experience with phantom vibration and ringing syndrome [22], and one item for quality of life measured on a 4-point Likert scale. In contrast to passive data, self-reported data from the questionnaire are called “active data” throughout this article.

Defining User Retention

Considering the intermittent usage of this app and patterns of uploaded data, we herein define that user retention of this app stopped on the first date of a 28-day period without any data being uploaded. For every smartphone, user retention was measured separately for passive and active data.

Statistical Analysis

In this study, the user retention length and indicator of whether an individual uninstalled the app were treated as a survival event. The Kaplan-Meier estimator, a standard nonparametric statistic used to estimate survival function of time-to-event data, was applied to measure the survival rate of user retention. A log-rank test was used to examine the difference in survival curves among different subpopulations or data collection modes (ie, active vs passive). A Cox proportional hazards model (Cox model) was used to assess the association between the hazard rate of user retention and two factors (smartphone users’ operating system and occupation). Occupations included resident physicians, visiting staff, medical students, nurses, and others. Comparing different occupations, we focused on resident physicians because their work hours are the longest among all occupations [22-27]. All analyses were performed using R software (version 3.4). The survival analysis was based on the package “survival” [28]. A *P* value less than .05 was considered statistically significant.

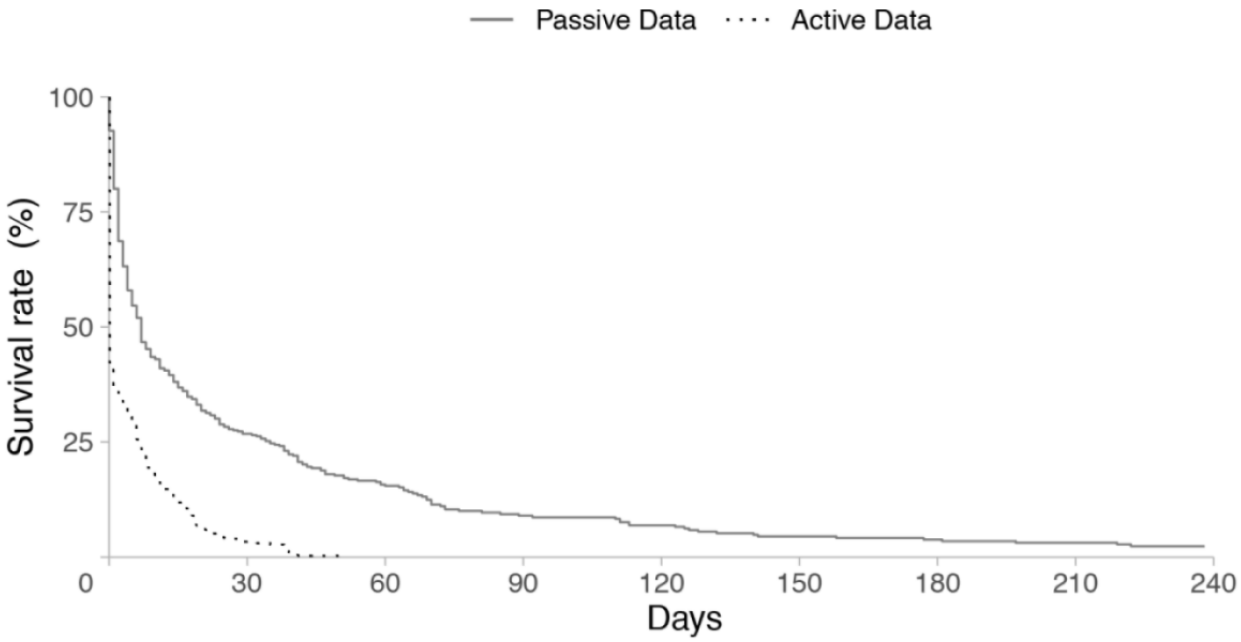
Patient and Public Involvement

No patients nor public members were involved in our work.

Results

We first compared user retention time between self-reported and GPS-defined work hours by analyzing active and passive data recorded using the “Staff Hours” app. Survival curves of user retention by passive and active data are illustrated in Figure 3. The survival rate of user retention for passive data was 46.7% after 1 week, whereas the rate for active data was 22.2%. After 1 month, the survival rate for passive data was 27.3%, and that for active data was 3.4%. The log-rank test results show that the survival time of passive data was significantly higher than that of active data (*P*=.011). This result is consistent after conditioning on the different operation systems, which is shown in Multimedia Appendix 1.

Figure 3. Survival curves of passive and active data.



The survival curves for the 2 operating systems (Android and iOS) for passive data are shown in Figure 4 and for active data are shown in Figure 5. For passive data, 1-week survival rates of user retention for Android and iOS were 69.2% and 36.2%,

respectively. After 1 month, the survival rate for Android reduced to 49.5%, whereas the rate for iOS reduced to 16.7%. For active data, 1-week survival rates of user retention for Android and iOS were 8.3% and 28.6%, respectively. After 1 month, the survival rate for Android reduced to 0.8%, and the survival rate for iOS reduced to 4.5%. The log-rank test revealed that patterns of user retention varied based on the operating system, with Android users maintaining a significantly higher survival rate of user retention for passive data across time (P=.015; Figure 4) than iOS users. By contrast, the patterns of user retention for active data showed no significant difference between Android and iOS (P=.700; Figure 5).

Figure 4. Survival curves of GPS-defined work hours between the operating systems.

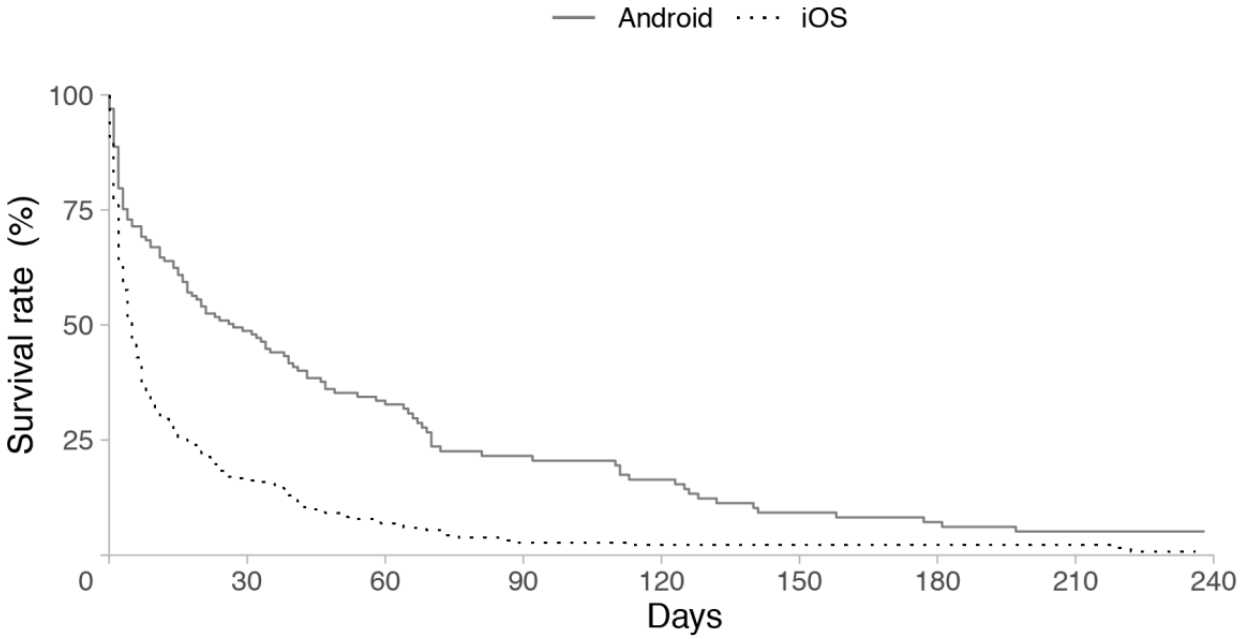
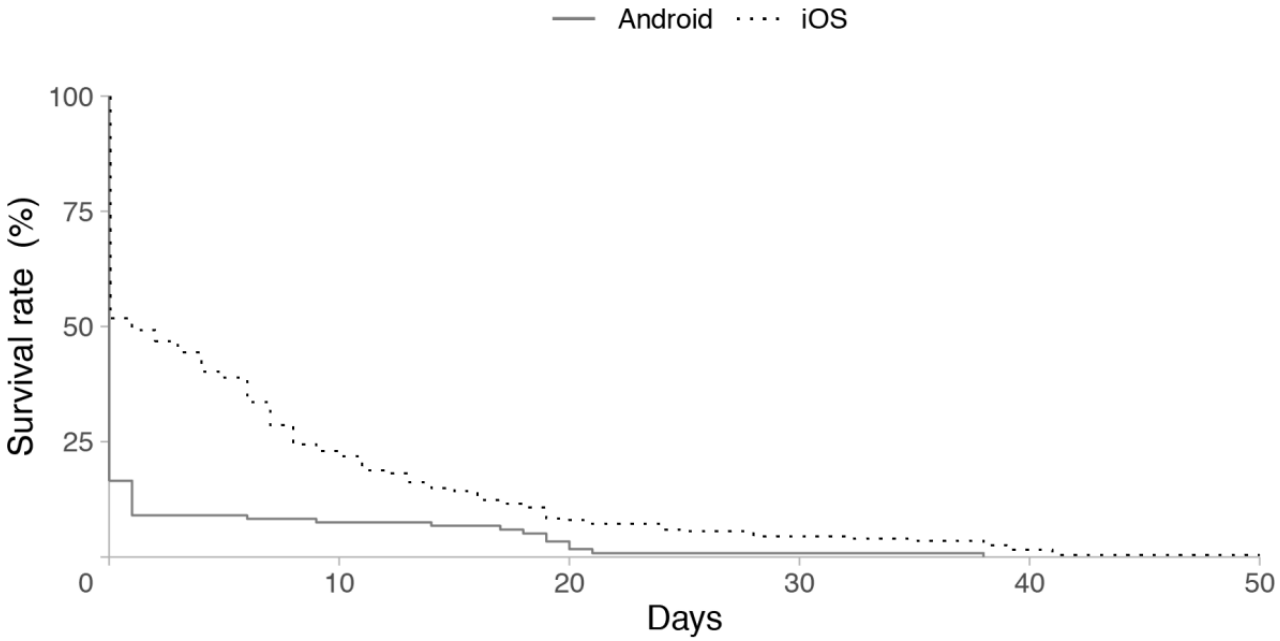


Figure 5. Survival curves of self-reports between the operating systems.



The results of the Cox model are shown in Table 1 and Multimedia Appendix 2. User retention for the collection of passive data via Android devices was higher than that for iOS. Resident physicians had higher user retention for the collection of active data than those with other occupations (P=.041), whereas no significant differences were observed between these two groups in passive data collection.

Table 1. Cox proportional hazards model.

Comparison	Passive data		Active data	
	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value
Operating system (iOS vs Android)	2.688 (1.846-3.913)	.012	1.707 (0.625-4.661)	.610
Occupation (resident physicians vs others)	0.711 (0.497-1.017)	.361	0.119 (0.044-0.322)	.041

Discussion

Principal Findings

Our results suggest that quantifying user retention of an app using survival analysis is feasible. The survival curves of user retention for passive and active data collection provide a reference for user retention of an app that was not actively promoted. This user retention obtained via survival analysis could be a useful indicator to monitor the effects of advertising campaigns or promotional activities for any app. Our data help to study the critical period of user retention. The survival rate of 46.7% for the collection of passive data was until the first week, and this suggests that more than half of the app users quit using the app or technical problems interrupted the passive data collection upon app installation. Helping first-time users to navigate this app and delivering frequently asked questions through a push notification during this critical period could be valuable, and this survival rate implies potential technical problems in the collection of passive data via this app. Previous studies have also demonstrated similar results — that user retention significantly decreases during the first week [29]. Among those who downloaded the "PTSD Coach" app (n=153,834), designed for American veterans to reduce post-traumatic stress disorder symptoms, 61.1% returned to use the app after the first day of installation. Over half of users (52.1%) continued to use the app or used it at least one time beyond the first week of download. A commonly observed pattern is the user downloading an app, quickly scanning it, and determining personal relevance. For users who may only open the app a single time, it is the only opportunity to capture their interest. Therefore, mobile health app developers have also suggested that health care professionals "need to very carefully manage the initial phases of somebody using this kind of technology and make sure they're well monitored" [30].

"Staff Hours" provides the first model to examine user retention for passive and active data collection simultaneously. Although a higher user retention for passive data collection over active data collection seems intuitive, we first quantified the survival rate of user retention and examined the factors associated with this survival rate. In addition, we established a user-centered evaluation of the mobile app's user retention based on an app-centered longevity evaluation [19]. The present analysis may be a useful metric of usability evaluation for a mobile health app, even a software medical device. Apart from the efficacy and safety evaluation in the traditional registration of new drugs or medical devices, the American Psychiatric Association proposed an app evaluation framework that emphasizes usability or engagement [31]. The engagement evaluation includes how many patients became stuck when using an app or found them difficult to use. This user-centered evaluation encourages app developers to involve patients or potential target users in the development of their health apps.

Our findings show that user retention for the collection of passive data via Android devices was higher than that with iOS. Platform-related differences in user retention resulted from the differences in the GPS data collection between Android and iOS. Specifically, "Staff Hours" will stop collecting data during a background refresh in iOS but not in Android. The fact that passive data collection consumed more electricity in iOS could explain the lower user retention in iOS for passive data. By contrast, no significant platform-related differences in user retention in the collection of active data were observed. Because active data collection differs from passive data collection, which was generated continuously from day-to-day human-machine interaction, user retention is not associated with operating systems. In summary, platform-related differences in user retention in the collection of passive data, but not of active data, imply that electricity consumption may be of a particular concern to smartphone users, given the challenges of developing apps for 2 distinct platforms.

Our results provide novel information regarding resident physicians' user retention in the collection of active data. This higher user retention suggests that resident physicians represent the right target audience for this app. "Staff Hours" was designed not only to automatically record work hours but also to survey staff's mental health during moments of long work hours. Resident physicians had extremely long work hours and mental health issues at the time. As a result, they were more motivated to continuously interact with this app. The PHQ-9 for depressive symptoms, phantom vibration, and ringing syndrome in this mental health survey is specifically designed for staff with excessive work hours such as resident physicians. A systematic review and meta-analysis showed a high prevalence rate of 28.8% of depression or depressive symptoms among resident physicians [32]. Our previous study also showed that more than 85% of medical interns with 86.7 work hours per week experienced phantom vibration and ringing syndrome, and these syndromes significantly reduced 2 weeks after their internship [27]. Such mental health surveys on smartphones are important because smartphone use may reduce bias in the form of the Hawthorne effect. This effect has been reported in a previous study to demonstrate that PHQ-9 depressive symptom scores recorded from the app were more sensitive in detecting suicidal behavior than the traditionally administered PHQ-9 [33]. Mental health and work hours are both important for resident physicians, but excessive work hours did not represent poor mental health. Our recent study showed that medical interns with an additional 10 work hours per week (ie, average additional 2 work hours per day) had a relatively small increase in depressive symptoms, with a PHQ-9 score of 0.13 [34].

Limitations

Several methodologic limitations should be noted when interpreting our findings. First, the rapid decrease in user retention in the collection of active data may be attributed to our definition of user retention, which stopped when no data were uploaded for more than 28 consecutive days. Furthermore, passive data and active data are collected from different sources (ie, GPS and self-report, respectively). Therefore, the difference in the user retention time is somehow affected by the characteristics of these resources. For example, sampling rates of active and passive data are unequal, and passive data are generated every 10 minutes, but active data, according to a mental health survey, are generated every month. Second, this definition of retention did not include users' return to using the app after quitting for more than 28 consecutive days. Third, our sample size was not large enough to identify more factors associated with user retention and lacked significant power. In addition, all of the individuals only had one smartphone, so it was not feasible to compare the difference in retention time between iOS and Android operation systems for the same individuals. Finally, the user retention might also be different in an app that is more attractive than "Staff Hours." In addition, the longer retention with our app on Android devices, compared with iOS devices, can be explained by the difference in electronic consumption, so the results might not be generalizable to other apps.

Conclusions

In conclusion, we demonstrated that passive data collected via Android devices had the best user retention with our app that records GPS-based work hours. As a pilot study in this field, our results provide new insights into quantifying the usability of a mobile app using survival analysis. We also determined that the first week upon installation is the critical period for the app's longevity. Analysis of user retention with additional apps is required to validate our methods.

Acknowledgments

We thank Ms. Ting-Wei Chiang and Ms. Hsiao-Han Lin for their excellent technical assistance. This study was supported by grants from the National Health Research Institutes of Taiwan (09A1-PHPP20-014) and Ministry of Science and Technology of Taiwan (MOST 107-2314-B-400-038-MY3 and MOST 108-2636-B-009-001). The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data and had final responsibility for the decision to submit for publication.

Authors' Contributions

YL, SL, and SC conceptualized the study. SC, PL, AT, YP, and CH performed the data curation and data analysis. All authors interpreted the data analysis and critically revised the manuscript. All authors had the opportunity to review the final manuscript and provided their permission to publish the manuscript. All authors agree to take responsibility for the work. YL is guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survival curves of GPS-defined work hours (passive data) vs self-reported work hours (active data) for each operating system.
[DOCX File , 25 KB - mhealth_v8i11e16309_app1.docx]

Multimedia Appendix 2
Cox proportional hazards model adjusted by age and gender.
[DOCX File , 13 KB - mhealth_v8i11e16309_app2.docx]

References

1. 5.2 billion mobile broadband subscriptions. Ericsson. 2018 Feb 20. URL: https://www.ericsson.com/en/news/2018/2/5.2-billion-mobile-broadband-subscriptions [accessed 2019-07-17]
2. Insel TR. Digital Phenotyping: Technology for a New Science of Behavior. JAMA 2017 Oct 03;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]
3. The Mobile Economy 2019. GSMA. 2019. URL: https://www.gsma.com/r/mobileeconomy/ [accessed 2019-07-17]
4. Blumenstock J, Cadamuro G, On R. Predicting poverty and wealth from mobile phone metadata. Science 2015 Nov 27;350(6264):1073-1076. [doi: 10.1126/science.aac4420] [Medline: 26612950]
5. 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]

6. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. Using smartphone technology to monitor physical activity in the 10,000 Steps program: a matched case-control trial. J Med Internet Res 2012 Apr 20;14(2):e55 [FREE Full text] [doi: 10.2196/jmir.1950] [Medline: 22522112]
7. Clement I, Lorenz A, Ulm B, Plidschun A, Huber S. Implementing Systematically Collected User Feedback to Increase User Retention in a Mobile App for Self-Management of Low Back Pain: Retrospective Cohort Study. JMIR Mhealth Uhealth 2018 Jun 06;6(6):e10422 [FREE Full text] [doi: 10.2196/10422] [Medline: 29875088]
8. Baumel A, Muench F, Edan S, Kane JM. Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
9. Torous J, Kiang MV, Lorme J, Onnela J. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. JMIR Ment Health 2016 May 05;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
10. Landrigan CP, Barger LK, Cade BE, Ayas NT, Czeisler CA. Interns' compliance with accreditation council for graduate medical education work-hour limits. JAMA 2006 Sep 06;296(9):1063-1070. [doi: 10.1001/jama.296.9.1063] [Medline: 16954485]
11. Landrigan CP, Rahman SA, Sullivan JP, Vittinghoff E, Barger LK, Sanderson AL, et al. Effect on Patient Safety of a Resident Physician Schedule without 24-Hour Shifts. N Engl J Med 2020 Jun 25;382(26):2514-2523. [doi: 10.1056/nejmoa1900669]
12. Chien J, Chen C, Lin S, Lin S, Lin Y. Cardiac Autonomic Modulation during on-Call Duty under Working Hours Restriction. Int J Environ Res Public Health 2020 Feb 10;17(3):1118 [FREE Full text] [doi: 10.3390/ijerph17031118] [Medline: 32050580]
13. Pan Y, Lin H, Chiu Y, Lin S, Lin Y. Temporal Stability of Smartphone Use Data: Determining Fundamental Time Unit and Independent Cycle. JMIR Mhealth Uhealth 2019 Mar 26;7(3):e12171 [FREE Full text] [doi: 10.2196/12171] [Medline: 30912751]
14. Lin Y, Wong B, Lin S, Chiu Y, Pan Y, Lee Y. Development of a mobile application (App) to delineate "digital chronotype" and the effects of delayed chronotype by bedtime smartphone use. J Psychiatr Res 2019 Mar;110:9-15. [doi: 10.1016/j.jpsychires.2018.12.012] [Medline: 30611008]
15. Lin Y, Lin P, Chiang C, Lee Y, Yang CCH, Kuo TBJ, et al. Incorporation of Mobile Application (App) Measures Into the Diagnosis of Smartphone Addiction. J. Clin. Psychiatry 2017 Aug 23;78(7):866-872. [doi: 10.4088/jcp.15m10310]
16. Lin Y, Lin Y, Lin S, Lee Y, Lin P, Chiang C, et al. To use or not to use? Compulsive behavior and its role in smartphone addiction. Transl Psychiatry 2017 Feb 14;7(2):e1030-e1030 [FREE Full text] [doi: 10.1038/tp.2017.1] [Medline: 28195570]
17. Lin Y, Lin Y, Lee Y, Lin P, Lin S, Chang L, et al. Time distortion associated with smartphone addiction: Identifying smartphone addiction via a mobile application (App). J Psychiatr Res 2015 Jun;65:139-145. [doi: 10.1016/j.jpsychires.2015.04.003] [Medline: 25935253]
18. Lin Y, Wong B, Pan Y, Chiu Y, Lee Y. Validation of the Mobile App-Recorded Circadian Rhythm by a Digital Footprint. JMIR Mhealth Uhealth 2019 May 16;7(5):e13421 [FREE Full text] [doi: 10.2196/13421] [Medline: 31099340]
19. Larsen ME, Nicholas J, Christensen H. Quantifying App Store Dynamics: Longitudinal Tracking of Mental Health Apps. JMIR Mhealth Uhealth 2016 Aug 09;4(3):e96 [FREE Full text] [doi: 10.2196/mhealth.6020] [Medline: 27507641]
20. Singh K, Drouin K, Newmark LP, Lee J, Faxvaag A, Rozenblum R, et al. Many Mobile Health Apps Target High-Need, High-Cost Populations, But Gaps Remain. Health Aff (Millwood) 2016 Dec 01;35(12):2310-2318. [doi: 10.1377/hlthaff.2016.0578] [Medline: 27920321]
21. Liu S, Yeh Z, Huang H, Sun F, Tjung J, Hwang L, et al. Validation of Patient Health Questionnaire for depression screening among primary care patients in Taiwan. Compr Psychiatry 2011 Jan;52(1):96-101. [doi: 10.1016/j.comppsych.2010.04.013] [Medline: 21111406]
22. Lin Y, Chen C, Li P, Lin S. A dimensional approach to the phantom vibration and ringing syndrome during medical internship. J Psychiatr Res 2013 Sep;47(9):1254-1258. [doi: 10.1016/j.jpsychires.2013.05.023] [Medline: 23786911]
23. Chen C, Lin S, Li P, Huang W, Lin Y. The Role of the Harm Avoidance Personality in Depression and Anxiety During the Medical Internship. Medicine 2015;94(2):e389. [doi: 10.1097/md.0000000000000389]
24. Lin Y, Chen C, Lin S, Liu C, Weng W, Kuo TBJ, et al. Gender differences in cardiac autonomic modulation during medical internship. Psychophysiology 2013 Jun 20;50(6):521-527. [doi: 10.1111/psyp.12038] [Medline: 23517383]
25. Lin Y, Ho Y, Lin S, Yeh Y, Liu C, Kuo TBJ, et al. On-call duty effects on sleep-state physiological stability in male medical interns. PLoS One 2013 Jun 4;8(6):e65072 [FREE Full text] [doi: 10.1371/journal.pone.0065072] [Medline: 23750232]
26. Lin Y, Kuo TBJ, Ho Y, Lin S, Liu C, Yang CCH. Physiological and psychological impacts on male medical interns during on-call duty. Stress 2012 Jan 19;15(1):21-30. [doi: 10.3109/10253890.2011.572208] [Medline: 21682655]
27. Lin Y, Lin S, Li P, Huang W, Chen C. Prevalent hallucinations during medical internships: phantom vibration and ringing syndromes. PLoS One 2013 Jun 10;8(6):e65152 [FREE Full text] [doi: 10.1371/journal.pone.0065152] [Medline: 23762302]
28. Therneau TM, Lumley T, Atkinson E, Crowson C. survival: Survival Analysis. CRAN.R project. 2020 Sep 28. URL: http://cran.r-project.org/web/packages/survival/ [accessed 2020-10-04]

29. Owen JE, Jaworski BK, Kuhn E, Makin-Byrd KN, Ramsey KM, Hoffman JE. mHealth in the Wild: Using Novel Data to Examine the Reach, Use, and Impact of PTSD Coach. JMIR Ment Health 2015 Mar 25;2(1):e7 [FREE Full text] [doi: 10.2196/mental.3935] [Medline: 26543913]
30. Anthes E. Mental health: There's an app for that. Nature 2016 Apr 07;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
31. Torous JB, Chan SR, Gipson SYT, Kim JW, Nguyen T, Luo J, et al. A Hierarchical Framework for Evaluation and Informed Decision Making Regarding Smartphone Apps for Clinical Care. Psychiatr Serv 2018 May 01;69(5):498-500. [doi: 10.1176/appi.ps.201700423] [Medline: 29446337]
32. Mata DA, Ramos MA, Bansal N, Khan R, Guille C, Di Angelantonio E, et al. Prevalence of Depression and Depressive Symptoms Among Resident Physicians: A Systematic Review and Meta-analysis. JAMA 2015 Dec 08;314(22):2373-2383 [FREE Full text] [doi: 10.1001/jama.2015.15845] [Medline: 26647259]
33. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a Personal Smartphone Custom App to Assess the Patient Health Questionnaire-9 (PHQ-9) Depressive Symptoms in Patients With Major Depressive Disorder. JMIR Ment Health 2015 Mar 24;2(1):e8 [FREE Full text] [doi: 10.2196/mental.3889] [Medline: 26543914]
34. Lin Y, Chen H, Tsai S, Chang L, Chen P. A prospective study of the factors associated with life quality during medical internship. PLoS One 2019 Aug 12;14(8):e0220608 [FREE Full text] [doi: 10.1371/journal.pone.0220608] [Medline: 31404080]

Abbreviations

PHQ-9: Patient Health Questionnaire

Edited by G Eysenbach; submitted 19.09.19; peer-reviewed by LA Lee, T Wu; comments to author 15.06.20; revised version received 23.07.20; accepted 15.09.20; published 26.11.20.

Please cite as:
Lin YH, Chen SY, Lin PH, Tai AS, Pan YC, Hsieh CE, Lin SH
Assessing User Retention of a Mobile App: Survival Analysis
JMIR Mhealth Uhealth 2020;8(11):e16309
URL: http://mhealth.jmir.org/2020/11/e16309/
doi:10.2196/16309
PMID:33242023

Original Paper

Sleep Tracking of a Commercially Available Smart Ring and Smartwatch Against Medical-Grade Actigraphy in Everyday Settings: Instrument Validation Study

Milad Asgari Mehrabadi[1], BSc, MSc; Iman Azimi[2], PhD; Fatemeh Sarhaddi[2], BSc, MSc; Anna Axelin[3,4], PhD; Hannakaisa Niela-Vilén[3], PhD; Saana Myllyntausta[5,6,7], PhD; Sari Stenholm[5,7], PhD; Nikil Dutt[1,8], PhD; Pasi Liljeberg[2], PhD; Amir M Rahmani[1,8,9], PhD, MBA

[1]Department of Electrical Engineering and Computer Science, University of California Irvine, Irvine, CA, United States
[2]Department of Computing, University of Turku, Turku, Finland
[3]Department of Nursing Science, University of Turku, Turku, Finland
[4]Department of Obstetrics and Gynaecology, Turku University Hospital, Turku, Finland
[5]Department of Public Health, University of Turku and Turku University Hospital, Turku, Finland
[6]School of Educational Sciences and Psychology, University of Eastern Finland, Joensuu, Kuopio, Finland
[7]Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland
[8]Department of Computer Science, University of California Irvine, Irvine, CA, United States
[9]School of Nursing, University of California Irvine, Irvine, CA, United States

Corresponding Author:
Milad Asgari Mehrabadi, BSc, MSc
Department of Electrical Engineering and Computer Science
University of California Irvine
Berk Hall, 1st Floor
Irvine, CA
United States
Phone: 1 949 506 8187
Email: masgarim@uci.edu

Abstract

Background: Assessment of sleep quality is essential to address poor sleep quality and understand changes. Owing to the advances in the Internet of Things and wearable technologies, sleep monitoring under free-living conditions has become feasible and practicable. Smart rings and smartwatches can be employed to perform mid- or long-term home-based sleep monitoring. However, the validity of such wearables should be investigated in terms of sleep parameters. Sleep validation studies are mostly limited to short-term laboratory tests; there is a need for a study to assess the sleep attributes of wearables in everyday settings, where users engage in their daily routines.

Objective: This study aims to evaluate the sleep parameters of the Oura ring along with the Samsung Gear Sport watch in comparison with a medically approved actigraphy device in a midterm everyday setting, where users engage in their daily routines.

Methods: We conducted home-based sleep monitoring in which the sleep parameters of 45 healthy individuals (23 women and 22 men) were tracked for 7 days. Total sleep time (TST), sleep efficiency (SE), and wake after sleep onset (WASO) of the ring and watch were assessed using paired *t* tests, Bland-Altman plots, and Pearson correlation. The parameters were also investigated considering the gender of the participants as a dependent variable.

Results: We found significant correlations between the ring's and actigraphy's TST (r=0.86; P<.001), WASO (r=0.41; P<.001), and SE (r=0.47; P<.001). Comparing the watch with actigraphy showed a significant correlation in TST (r=0.59; P<.001). The mean differences in TST, WASO, and SE of the ring and actigraphy were within satisfactory ranges, although there were significant differences between the parameters (P<.001); TST and SE mean differences were also within satisfactory ranges for the watch, and the WASO was slightly higher than the range (31.27, SD 35.15). However, the mean differences of the parameters between the watch and actigraphy were considerably higher than those of the ring. The watch also showed a significant difference in TST (P<.001) between female and male groups.

Conclusions: In a sample population of healthy adults, the sleep parameters of both the Oura ring and Samsung watch have acceptable mean differences and indicate significant correlations with actigraphy, but the ring outperforms the watch in terms of the nonstaging sleep parameters.

(JMIR Mhealth Uhealth 2020;8(10):e20465) doi:10.2196/20465

KEYWORDS

sleep; smart ring; smartwatch; actigraphy; wearable technology

Introduction

Background

Sleep is a multifaceted and dynamic phenomenon that indicates individuals' overall health and well-being and is affected by a variety of factors such as behavioral habits, stress, and disorders [1,2]. Sleep disturbances are common across different population groups (eg, older people and pregnant women) and negatively impact body functions, including the cardiovascular and immune system and hormonal release [3,4]. Such sleep problems need to be investigated thoroughly to reduce the associated health risks and complications. Monitoring sleep quality is a vital step in this regard when the individuals' sleep parameters are tracked [5].

Sleep quality assessment methods have been conventionally performed in clinical settings by monitoring users' biological signals and body movements. Polysomnography (PSG), the gold standard method used for sleep analysis, is enabled by the continuous monitoring of different cardiorespiratory and neurophysiological indicators [6]. Owing to PSG's complex and multichannel data collection, this method is limited to short-term hospital or laboratory-based monitoring. Actigraphy is another well-established method enabled by a 3D accelerometer that captures the movements of a limb to monitor sleep [7]. This method has been shown to be accurate enough compared with PSG in a healthy subject population [8-11], although the results might be inaccurate when the subjects are individuals with sleep disorders [7,11,12]. In addition, other studies conducted with large populations have shown an agreement between actigraphy and PSG in total sleep time (TST), wake after sleep onset (WASO), and sleep efficiency (SE) parameters [11,13]. On the other hand, some studies have considered the validity of actigraphy's sleep onset latency (SOL) compared with PSG [11,14] and showed that actigraphy consistently underestimated SOL in comparison with PSG. This method is more convenient than PSG because it allows users to wear the actigraphy device in everyday settings (ie, days to weeks), although conventional medical-grade actigraphy devices are still infeasible for long-term studies (ie, months to years) because of their size, design, and battery life issues.

Advancements in consumer wearable technology provide opportunities to extend sleep monitoring to mid- or long-term home-based health care applications using low-power, miniaturized, and fashionable wearables [15-17]. Wearable electronics and the Internet of Things–based systems are growing dramatically and are expected to revolutionize health care delivery and outcomes [18,19]. In particular, smart rings will most likely become popular in sleep studies. Longer battery life, elegant design, and sophisticated embedded sensors in such rings have enabled them to be used not only in clinical trials (instead of medical-grade actigraphy) but also in different population-based studies [20,21]. Such devices offer continuous data collection of body movements and vital signs in everyday settings. The data can be utilized to continuously monitor sleep disturbances of individuals for an extended period [22].

Sleep monitoring using consumer wearables such as wrist-worn activity trackers, smartwatches, and smart rings necessitate valid sleep data collection and data analysis to provide accurate sleep parameters. Various studies have investigated wrist bands in terms of sleep monitoring accuracy across different population groups. For example, the validation of sleep data of 7 different commercial activity trackers was assessed by conducting data collection for 2 days on healthy adults [23]. In other studies, the sleep estimation of Fitbit devices [24-26], Jawbone [27-29], and physical activity monitors [30] has been investigated against actigraphy, PSG, or both in overnight tests on healthy adolescents and individuals with obstructive sleep apnea. These studies focused on the sleep quality assessment of wearables by tracking a set of nonstaging sleep parameters, including TST, SOL, WASO, and SE [31-34]. Regarding smart ring validation, there is one study that has validated the Oura smart ring against PSG in an overnight laboratory setup [35]; however, there is no previous research in the literature validating a smart ring against actigraphy in the mid- or long-term. Furthermore, these earlier validation studies are limited to laboratory settings and/or overnight (ie, single night) data collection. The effect of home-based health monitoring, where the users might be involved in different conditions and environments, is ignored in these validation studies. Therefore, the results obtained could be inaccurate for long-term and remote monitoring.

Objectives

In this paper, we aim to assess the validity of sleep data acquired by a smart ring, Oura, in comparison with a medically approved actigraphy device. We utilize the Oura ring as a compact and relatively small device with a user-friendly design. In addition, we assessed the Samsung Gear Sport smartwatch against actigraphy to compare the accuracy of Oura ring in the detection of different sleep attributes. In general, because watches and rings are worn in different parts of the subject's hand, they respond differently to signal logging disturbances, such as *motion artifacts*. The devices were tested in a 7-day monitoring study, approved by the ethical committee, where the sleep data of 45 healthy individuals were monitored. Participants were asked to use the devices 24 hours for 7 days and carry out their daily routines as usual. We compared TST, SOL, WASO, and SE obtained from the Oura ring, Samsung watch, and ActiGraph. The parameters obtained by the 2 consumer-grade wearables (ie, the ring and the watch) were evaluated with the sleep

parameters extracted from actigraphy using paired *t* tests, Bland-Altman [36] plots, and Pearson correlation. The parameters were investigated considering the gender of the participants as a dependent variable. Finally, we conclude the paper with a discussion of our obtained results and the validity of sleep data of the wearables in everyday settings.

Methods

Participants and Recruitment

Recruitment was performed in southern Finland from July to August 2019. In earlier validation studies of commercial devices, the sample sizes varied between 20 and 40. Therefore, we aimed at a target sample of 40 people. The recruitment started with convenience sampling by personally contacting a few students and staff members of the University of Turku. Afterward, snowball sampling was used until the target sample size was reached; 6 additional participants were enrolled because of expected missing data. We aimed for variation among participants by age, weight, physical activity, education, and lifestyle as related to sleep and stress levels.

A sample of healthy individuals between 18 and 55 years of age was enrolled. Potential participants were excluded if they had (1) a diagnosed cardiovascular disease, (2) restrictions regarding physical activity, (3) symptoms of an illness at the time of recruitment (ie, flu symptoms including sore throat, runny nose, cough, and fever), or (4) any restrictions on using the devices at work. In a face-to-face meeting with the interested individuals, researchers described the purpose of the study and the wearable devices. They were asked to wear the Gear Sport smartwatch, Oura ring, and ActiGraph wristband for 1 week in their normal daily life. Each participant provided written informed consent. Altogether, 46 participants, including 23 women and 23 men, participated in the study. A participant (male) was excluded from the analysis because he did not wear the actigraphy device. Therefore, the final sample size was 45 (23 women, 22 men). Table 1 shows the participants' background information. The table includes 42 participants, as the background information of the 3 participants is missing.

Table 1. Participants' background information.

Characteristics	Values
Age (years), mean (SD)	
Women	31.5 (6.6)
Men	33 (6)
BMI, mean (SD)	
Women	24.4 (5.6)
Men	25.5 (2.9)
Expected sleep (daily hours), mean (SD)	
Women	7.35 (1.00)
Men	7.17 (1.05)
Physical activity, n (%)	
Almost daily	12 (27)
Once a week	9 (20)
>Once a week	21 (47)
Working status, n (%)	
Working	32 (71)
Unemployed	1 (2)
Student	8 (18)
Other	1 (2)

Ethics

The study was conducted according to the ethical principles based on the Declaration of Helsinki and the Finnish Medical Research Act (#488/1999). The study protocol received a favorable statement from the ethics committee (University of Turku, Ethics committee for Human Sciences, Statement #44/2019). The participants were informed about the study, both orally and in writing, before obtaining their consent. Participation was voluntary, and all participants had the right to withdraw from the study at any time and without giving any reason. To compensate for the time used for the study, each participant received a €20 (US $23) gift card to the grocery store at the end of the monitoring period when returning the devices.

Data Collection

Our data collection for 1 week included 4 approaches for monitoring participants' sleep. We utilized 3 devices (ie, 2 wearable and 1 actigraphy device) to continuously capture sleep

data and a self-report form by which subjective measures were collected. Samsung Gear and ActiGraph were worn in the wrist, and the Oura ring was worn in one of the fingers of the nondominant hand; thus, all 3 devices were on the same hand. The participants completed a short background questionnaire at the meetings. They were also asked to report their sleep times, such as bedtime, waking up time, and naps, during the 7-day study period via a structured self-report (ie, daily log) form. They were also asked to report other events during the study, such as device removal from the wrist or if specific events occurred (eg, visiting a hospital because of a disease). The self-report data were used to interpret the actigraphy data and mitigate possible errors; such a correction was necessary for this study because the actigraphy was selected as the baseline sleep monitoring method. In addition to the verbal instructions, participants were given a written guideline for using the devices.

The Oura ring [37] was the first wearable device investigated in this study. The Oura ring is a commercial sleep tracker device that uses acceleration and gyroscope data, photoplethysmogram (PPG) signal, and body temperature to estimate sleep parameters, heart rate variability, respiratory rate, and intensity of physical activity. The ring is lightweight (4-6 g) and easy to use. It also has an acceptable battery life, that is, the battery lasts up to 1 week in regular use. The ring is connected to the Android or iOS Oura mobile app via Bluetooth. The data are automatically sent to the mobile app and transferred to the cloud server. The data can be accessed from the mobile app or the cloud server. In this study, we extracted the sleep data of participants from the Oura cloud.

In addition to the Oura ring, we used the Samsung Gear Sport watch [38], which is an open-source smartwatch that enables remote health monitoring. The watch includes a PPG sensor and an inertial measurement unit through which PPG signal, acceleration, and gyroscope data can be collected continuously. The data are processed to extract various variables, including heart rate, sleep duration, and step counts. The Gear Sport watch runs open-source Tizen operating system, enabling customized data collection. In this study, we programmed the watch to collect sleep parameters, PPG data, and hand movement data during the monitoring. The PPG and hand movement data were utilized to validate the sleep events (detailed in the *Data Analysis* section). Moreover, we also developed an app for the watch to send the collected data to our server via Wi-Fi.

For actigraphy, we used the wGT3X-BT device by ActiGraph. The wGT3X is a noncommercial triaxial accelerometer that measures the wrist's acceleration in 3 orthogonal axes at 80 Hz. This device is waterproof, and its battery life is approximately 3 weeks. The device does not provide any feedback to the participants about their activity or sleep. The acceleration data collected by the device were utilized to obtain the estimates of sleep parameters.

Data Analysis

Data analysis included the sleep parameter extraction from the collected data and the statistical analysis leveraged to evaluate the ring and watch.

Actigraphy

Raw data from the actigraphy device were downloaded to a computer and converted into 60-second epochs using the ActiLife software (version 6.13) [39] provided by the manufacturer (ActiGraph). We used the Cole-Kripke algorithm [40] to define each epoch as sleep or wake. This algorithm was selected because it has been validated in the adult population using wrist-worn accelerometers. The ActiGraph algorithm that is available in the ActiLife software was then used to detect the sleep periods and estimate sleep attributes. Using the Troiano wear time validation algorithm [41], the auto sleep period detection algorithm detects nonwear bouts, ignores nonwear periods greater than a day, and nonwear periods that have almost all zeros (5 or more epochs of nonzeros). The nonwear periods that remain are defined as sleep time. Sleep data were systematically checked, cleaned, and sleep periods that did not represent true sleep times were deleted. These deletions included sleep periods with nonwear time during evenings or mornings that the algorithm had incorrectly scored as sleep, daytime sleep periods, and sleep periods outside the actual measurement week.

Wearables

We used the application programming interface provided by the Oura ring and the Samsung watch to extract different semistructured data for our analyses. The Oura ring provides JavaScript Object Notation files, including the sleep parameters per night. The 3 main types of sleep parameters provided by the ring are (1) parameters related to different levels of sleep and nonstaging sleep, including the start and end of sleep, the number of awakenings, total awakening time, and sleep onset, (2) scores to measure the quality of sleep in different stages, and (3) average heart rate for every 5 min during sleep. In this study, we only investigated the nonstaging sleep parameters because of the limitation of the baseline actigraphy method.

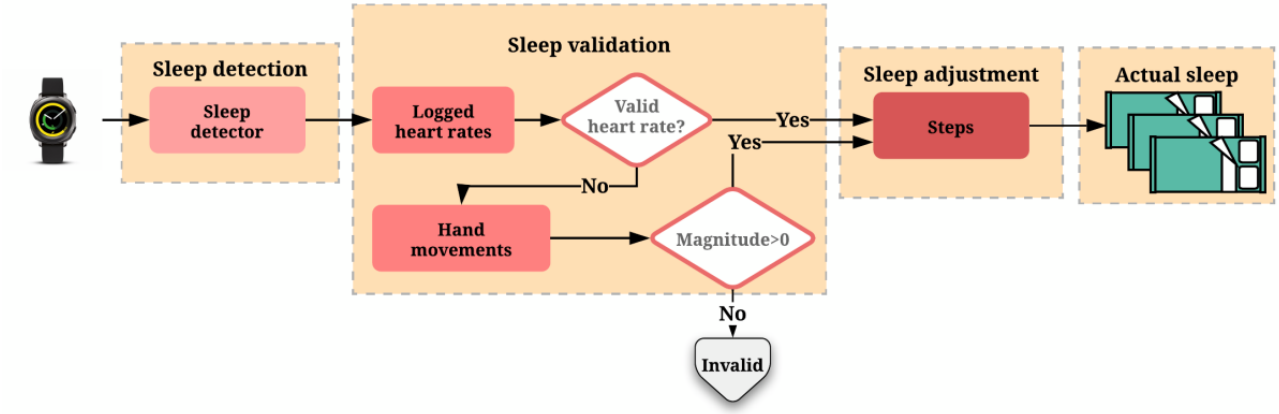
In contrast, the Gear Sport watch provides a data record when the user's status changes; for example, the status changes from wake to sleep. We used these records to extract sleep events per night and validated the sleep events using the heart rate and hand movement data collected by the watch. Validation was performed to prevent the misdetection of sleep events owing to not wearing the device. For example, the watch was not used (no movement) for 1 hour, but a sleep event was detected by mistake. In this regard, we recorded a window of 30-second PPG signal when a sleep event started and ended. The sleep event was considered valid if valid heart rate values were detected from the PPG signals. In addition, we considered the hand movement magnitude for validation if the PPG signal was invalid because of practical issues. Finally, we cross-checked the sleep events with the step count data (reported by the watch) and corrected or discarded the sleep events if there was no match between the data.

It should be noted that the watch could not detect a few sleep events because of technical and practical issues during the monitoring. For example, the sleep event was missed because the watch's turn-off button was pressed accidentally during the night. This issue mostly occurred during the monitoring, as the watch and actigraphy were worn on the same hand close to each other. As the watch could not record the sleep events, we

removed 21 nights of data out of 181 (21/181, 11.6%) of the watch for the sake of an unbiased comparison between the actigraphy and watch.

Using the actual valid sleep events, we calculated WASO, TST (in minutes), and SE (%) per night. As the watch does not provide SOL explicitly, we calculate such a feature based on the difference between the start of the actual sleep and the last time the subject had steps. A summary of the processing pipeline is illustrated in Figure 1.

Figure 1. Watch data processing pipeline.



Statistical Analysis

We report the mean, SD, and 95% CI of the sleep parameters collected by the Oura ring, Samsung watch, and ActiGraph. The difference between the ring (or the watch) and the ActiGraph was also computed using two-tailed paired *t* tests to test the null hypothesis. In our context, the null hypothesis is that the true mean difference between the two measurements is 0 [42]. Due to the interest in observing the paired differences between values reported by ring (or the watch) and ActiGraph (baseline), the paired *t* test was utilized. In addition, we used the Bland-Altman plot to illustrate and estimate the agreement between the devices. These methods provided mean differences (bias) and SD of the differences between the ring (or the watch) and the actigraphy, lower and upper agreement limits, and 95% CI of the mean differences. The sign of mean differences indicates underestimation or overestimation of the ring (or the watch) compared with the actigraphy: a negative bias shows an overestimation, whereas a positive bias indicates an underestimation.

The satisfactory difference between the ring (or the watch) and the actigraphy data was selected as ≤30 min for TST and WASO and <5% for SE, similar to other studies in the literature [27,35,43]. We investigated the ratio of the samples within these satisfactory ranges. Moreover, we also investigated gender as a dependent variable in the validity of sleep parameters using *t* tests, considering the mean differences between the ring (or the watch) and the actigraphy.

Finally, to analyze the linear relationship between actigraphy and the ring (or the watch) corresponding sleep measurements, we performed Pearson correlation tests on pairwise sleep attributes of the actigraphy and the ring (or the watch).

Results

Study Population

A total of 45 subjects (23 women and 22 men) participated in this study. The subjects were 33.1 years old, on average, with an SD of 6.4 years. In total, we recorded 284 valid available days by actigraphy; however, after matching the corresponding available days of the ring (or the watch), we had fewer valid days for the analysis.

As discussed in the *Methods* section, in this study, we exploited 4 different sleep attributes. Although the results regarding SOL are not conclusive (because SOL of the actigraphy is unreliable [14]), for the sake of comparison, we report such results in addition to the other sleep parameters in this section.

Comparisons Between Ring and Equivalent Actigraphy Sleep Measures

To validate the Oura ring against actigraphy, we matched the available dates of the ring with the corresponding dates of actigraphy. In total, for all the participants, sleep data of 266 days (ie, 5.91, SD 1.32 days per subject) were included in the analysis.

The mean, SD, and 95% CI of the extracted sleep parameters are presented in Table 2. The table also shows the paired *t* test values of these parameters with their corresponding *P* values. Bland-Altman plots were used to show the agreements between the 2 measures. Figure 2 depicts the agreement between the ring and actigraphy for the TST, WASO, and SE. The bias and lower and upper agreement limits for these parameters are also summarized in Table 3.

Table 2. Mean, SD, 95% CI, and paired t test results for the actigraphy and the Oura ring sleep parameters in a sample of 45 healthy adults.

Parameter	Mean (SD)	95% CI	*t* value (*df*)	*P* value
Total sleep time (min)				
t test	N/A[a]	N/A	−6.26 (265)	<.001
Actigraphy	419.04 (78.31)	409.59-428.5	N/A	N/A
Oura ring	434.31 (72.14)	425.6-443.02	N/A	N/A
Sleep efficiency (%)				
t test	N/A	N/A	3.69 (265)	<.001
Actigraphy	90.47 (5.1)	89.86-91.09	N/A	N/A
Oura ring	89.13 (6.28)	88.38-89.89	N/A	N/A
Wake after sleep onset (min)				
t test	N/A	N/A	10.03 (265)	<.001
Actigraphy	43.57 (27.28)	40.28-46.86	N/A	N/A
Oura Ring	26.17 (24.98)	23.15-29.18	N/A	N/A

[a]N/A: not applicable.

Figure 2. Bland-Altman plots for total sleep time, sleep efficiency, and wake after sleep onset gathered by the Oura ring and the actigraphy device. Subjects' actigraphy minus Oura ring discrepancies on sleep parameters (y-axis) are plotted compared with actigraphy (x-axis). Biases, upper, and lower agreement limits are marked. In addition, the satisfactory ranges are plotted as the dashed lines. SE: sleep efficiency; TST: total sleep time; WASO: wake after sleep onset.

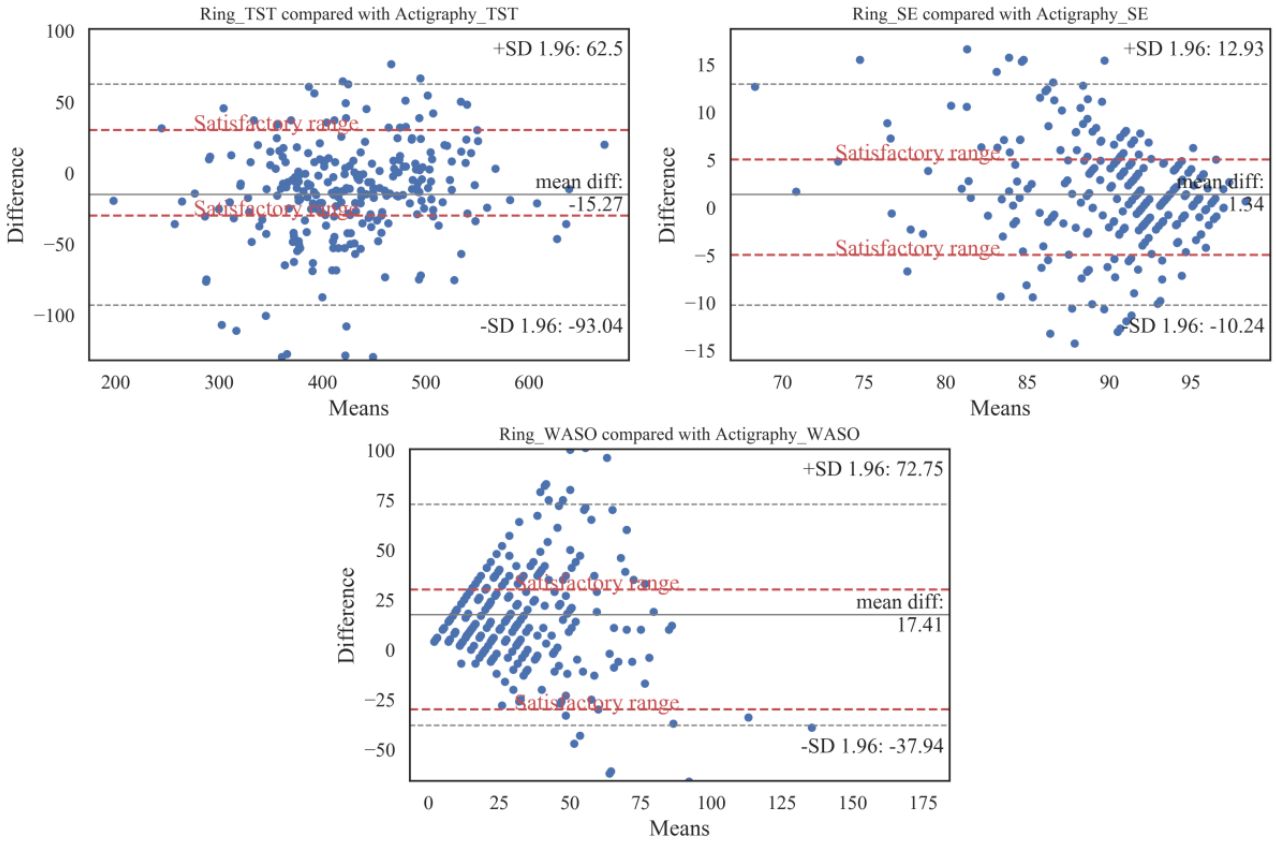


Table 3. Bias and agreement limits based on Bland-Altman plots for the actigraphy and the Oura ring.

Parameter	Mean difference (SD)	Lower and upper agreement limits
Total sleep time (min)	−15.27 (39.68)	−93.04, 62.5
Sleep efficiency (%)	1.34 (5.91)	−10.24, 12.93
Wake after sleep onset (min)	17.41 (28.24)	−37.94, 72.75

As shown in Table 2, the ring significantly overestimated the actigraphy (t_{265}=−6.26; *P*<.001) in the estimation of TST. On the basis of Figure 2, this overestimation in TST is, on average, 15.27 (SD 39.68) min (95% CI −20.07 to −10.47). Of 266 total samples, 14 fell outside the agreement range (lower limit −93.04 min, upper limit 62.50 min). The mean difference of TST between the actigraphy and ring fell within the satisfactory range, and 65.0% (173/266) of the data samples followed the satisfactory range condition.

On the other hand, in terms of WASO, the Oura ring significantly underestimated (t_{265}=10.03; *P*<.001) the actigraphy by, on average, 17.41 min (95% CI 13.99 to 20.82). Out of 266 samples, 17 fell outside the agreement limits (lower limit −37.94 min, upper limit 72.75 min). In terms of the satisfactory range, the mean difference fell within the range and covered 69.9% (186/266) of the total samples.

In addition, the Oura ring underestimated SE compared with the actigraphy by 1.34% on average (95% CI 0.63 to 2.06). This underestimation was significant, as shown in Table 2(*t_{265}*=3.69; *P*<.001). The mean difference in SE between the Oura ring and the actigraphy fell within the satisfactory range (<5%), along with 65.8% (175/266) of samples (including 44 out of 45 subjects). Moreover, 18 samples fell outside the agreement limits (lower limit −10.24%, upper limit 12.93%).

Comparisons Between Watch and Equivalent Actigraphy Sleep Measures

Similar to the ring validation, we considered the available dates for the actigraphy with corresponding data collected by the Samsung watch. As mentioned in the *Wearables* section, we removed the technically invalid watch data that occurred because of practical issues during the monitoring. Therefore, there were fewer sleep data from the watch than the other devices. After the matching procedure and invalid data removal, the number of subjects for the watch validation was 35 (19 men and 16 women), with 134 data samples (3.82, SD 1.50 days per subject). Table 4 summarizes the mean, SD, and 95% CI of the Samsung watch and the actigraphy with the corresponding available dates for different sleep parameters.

Table 4. Mean, SD, 95% CI, and paired t test results for the actigraphy and the Samsung watch sleep parameters in a sample of 35 healthy adults.

Parameter	Mean (SD)	95% CI	*t* value (*df*)	*P* value
Total sleep time (min)				
t test	N/A[a]	N/A	−3.54 (133)	<.001
Actigraphy	409.29 (81.43)	395.38-423.21	N/A	N/A
Samsung watch	431.81 (82.21)	417.76-445.85	N/A	N/A
Sleep efficiency (%)				
t test	N/A	N/A	−6.49 (133)	<.001
Actigraphy	90.40 (5.05)	89.54-91.26	N/A	N/A
Samsung watch	94.84 (7.03)	93.64-96.04	N/A	N/A
Wake after sleep onset (min)				
t test	N/A	N/A	10.26 (133)	<.001
Actigraphy	42.23 (23.43)	38.23-46.24	N/A	N/A
Samsung watch	10.96 (30.46)	5.76-16.17	N/A	N/A

[a]N/A: not applicable.

In addition, we performed paired *t* tests for the sleep parameters of the 2 devices. The results are shown in Table 4. As shown in this table, the *t* test values for all considered sleep parameters were statistically significant (*P*<.001). The positive and negative sign of the *t* value denotes the underestimation and overestimation of actigraphy by the watch, respectively. Bland-Altman plots showing TST, WASO, and SE agreements between the actigraphy and the watch are also illustrated in Figure 3. Moreover, bias and lower and upper agreement limits of sleep parameter outcomes by the actigraphy and the watch are summarized in Table 5.

Figure 3. Bland-Altman plots for total sleep time, sleep efficiency, and wake after sleep onset gathered by the Samsung watch and the actigraphy device. Subjects' actigraphy minus Samsung watch discrepancies on sleep parameters (y-axis) are plotted compared with actigraphy (x-axis). Biases, upper, and lower agreement limits are marked. In addition, the satisfactory ranges are plotted as the dashed lines. SE: sleep efficiency; TST: total sleep time; WASO: wake after sleep onset.

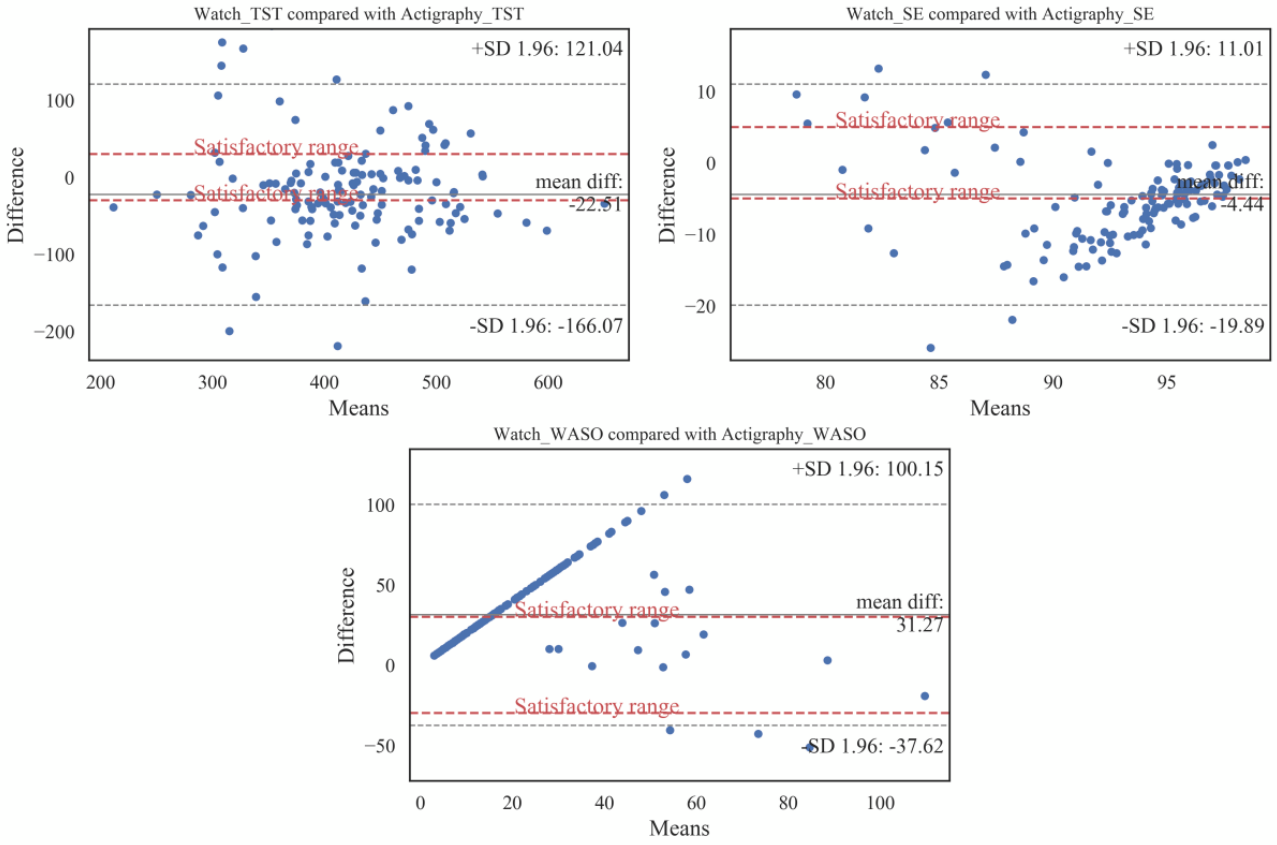


Table 5. Bias and agreement limits based on Bland-Altman plots for the actigraphy and the Samsung watch.

Parameter	Mean difference (SD)	Lower and upper agreement limits
Total sleep time (min)	−22.51 (73.24)	−166.07, 121.04
Sleep efficiency (%)	−4.44 (7.88)	−19.89, 11.01
Wake after sleep onset (min)	31.27 (35.15)	−37.62, 100.15

As shown in Figure 3, the watch overestimated the actigraphy in TST, on average, by 22.51 min (95% CI −35.08 to −9.95). Among the 134 samples, 9 were beyond the agreement limits (lower limit −166.07 min, upper limit 121.04 min). The mean difference of the actigraphy's and the watch's TST was within the satisfactory range; however, less than 50% (52/134, 38.8%) of the samples were within this satisfactory range.

In addition to TST, the Samsung watch overestimated SE by 4.44% (95% CI −5.79 to −3.09) compared with the actigraphy; 8 samples fell outside the agreement limits (lower limit −19.89%, upper limit 11.01%), with 42.5% (57/134) of the samples within the satisfactory range.

On the other hand, the watch underestimated WASO by 31.27 min on average (95% CI 25.24 to 37.3). Only 9 samples were outside of the agreement limits (lower limit −37.62 min, upper limit 100.15 min), and 45.5% (61/134) of the samples were within the satisfactory range.

Gender-Dependent Changes in the Mean Differences Between the Actigraphy and the Ring (or the Watch)

We also considered the gender of the participants to determine if the mean difference in sleep parameters differed between female and male groups. Table 6 shows the mean and SD of each sleep attribute of the actigraphy and the ring and the difference between these devices for male and female groups, separately.

The average of the mean difference between the TST of the actigraphy and the Oura ring did not differ between the male and female groups (t_{530}=0.99; P=.32). However, the mean differences of the other sleep parameters (ie, SE and WASO) were significant between female and male participants (P<.001 and P=.004).

Table 6. Mean, SD, and average mean differences (the actigraphy minus the Oura ring) for 23 women (141 samples) and 22 men (125 samples).

Parameter	Mean (SD)			*t* value (*df*)	*P* value
	Actigraphy	Oura ring	Differences		
Total sleep time (min)					
t test	N/A[a]	N/A	N/A	0.99 (530)	.32
Women	429.67 (70.25)	442.66 (64.67)	−12.98 (37.94)	N/A	N/A
Men	407.05 (85.21)	424.89 (78.94)	−17.84 (41.39)	N/A	N/A
Sleep efficiency (%)					
t test	N/A	N/A	N/A	−4.33 (530)	<.001
Women	90.64 (4.93)	90.73 (5.16)	−0.09 (5.86)	N/A	N/A
Men	90.29 (5.31)	87.33 (6.9)	2.96 (5.55)	N/A	N/A
Wake after sleep onset (min)					
t test	N/A	N/A	N/A	2.86 (530)	.004
Women	44.9 (30.08)	22.87 (20.7)	22.03 (29.19)	N/A	N/A
Men	42.07 (23.75)	29.88 (28.69)	12.19 (26.16)	N/A	N/A

[a]N/A: not applicable.

Similarly, we compared the mean differences of the sleep parameters between the actigraphy and the watch for the male and female groups. Table 7 summarizes such differences for each sleep parameter. As shown in Table 7, there was a significant difference between the mean differences of the male and female groups for TST (P<.001), SE (P=.01), and WASO (P=.01).

Table 7. Mean, SD, and average mean differences (the actigraphy minus the Samsung watch) for 16 women (65 samples) and 19 men (69 samples).

Parameter	Mean (SD)			*t* value (*df*)	*P* value
	Actigraphy	Samsung watch	Differences		
Total sleep time (min)					
t test	N/A[a]	N/A	N/A	3.48 (266)	<.001
Women	427.08 (73.76)	427.67 (74.76)	−0.59 (65.67)	N/A	N/A
Men	392.54 (85.22)	435.7 (89.04)	−43.16 (74.01)	N/A	N/A
Sleep efficiency (%)					
t test	N/A	N/A	N/A	2.39 (266)	.01
Women	90.82 (4.88)	93.6 (7.92)	−2.78 (8.04)	N/A	N/A
Men	90.0 (5.2)	96.01 (5.9)	−6.0 (7.4)	N/A	N/A
Wake after sleep onset (min)					
t test	N/A	N/A	N/A	−2.40 (266)	.01
Women	42.49 (24.33)	18.64 (39.75)	23.85 (42.54)	N/A	N/A
Men	41.99 (22.73)	3.73 (14.76)	38.26 (24.36)	N/A	N/A

[a]N/A: not applicable.

Correlations

We also investigated the possible linear relationship between the actigraphy and the ring (or the watch) data, using the Pearson correlation test. The correlation value (r) ranges from −1 to 1, where ±1 implies an exact linear relationship. The correlation values and their P values are shown in Table 8.

Table 8. Pearson correlation between the actigraphy, ring, and smartwatch with the corresponding *P* values for the considered sleep attributes.

Devices	Pearson correlation with the actigraphy, *r*					
	TST[a]	*P* value	SE[b]	*P* value	WASO[c]	*P* value
Oura ring	0.86	<.001	0.47	<.001	0.41	<.001
Samsung watch	0.59	<.001	0.17	.04	0.16	.06

[a]TST: total sleep time.

[b]SE: sleep efficiency.

[c]WASO: wake after sleep onset.

As shown in Table 8, comparing TST of actigraphy with TST of the ring and TST of the watch, we found a significantly high correlation between the actigraphy and the ring (r=0.86; P<.001). In contrast, the correlation between the actigraphy and the watch was r=0.59 (P<.001).

With regard to SE, there was a correlation between actigraphy and the ring (r=0.47; P<.001). In addition, the correlation between the actigraphy and the watch was acceptable (r=0.17; P=.04), but not as high as that of the ring.

For the WASO validation, there was a significant correlation between the actigraphy and the ring (r=0.41; P<.001). However, our analysis showed a nonsignificant correlation between WASO of the actigraphy and WASO of the watch (r=0.16; P=.06).

SOL Comparison Across Devices

As previously mentioned, SOL results were not conclusive since SOL of actigraphy is unreliable. We report SOL separately in the following: mean, SD, 95% CI, and paired *t* test results of the SOL for comparison between the actigraphy and the Oura ring (or Samsung watch) are presented in Tables 9 and 10. Bland-Altman plots showing the SOL agreements between the actigraphy and the ring (or the watch) are illustrated in Figures 4 and 5. Details of these plots are summarized in Tables 11 and 12.

Table 9. Mean, SD, 95% CI, and paired t test results for the actigraphy versus Oura ring estimates of sleep onset latency.

Parameter	Mean (SD)	95% CI	*t* value (*df*)	*P* value
Sleep onset latency (min)				
t test	N/A[a]	N/A	−13.01 (265)	<.001
Actigraphy	0.91 (1.37)	0.75-1.08	N/A	N/A
Oura ring	12.84 (14.92)	11.04-14.65	N/A	N/A

[a]N/A: not applicable.

Table 10. Mean, SD, 95% CI, and paired t test results for the actigraphy versus Samsung watch estimates of sleep onset latency.

Parameter	Mean (SD)	95% CI	*t* value (*df*)	*P* value
Sleep onset latency (min)				
t test	N/A[a]	N/A	−10.08 (133)	<.001
Actigraphy	0.99 (1.38)	0.75-1.22	N/A	N/A
Samsung watch	13.79 (14.86)	11.25-16.33	N/A	N/A

[a]N/A: not applicable.

Figure 4. Bland-Altman plot for sleep onset latency estimated by the Oura ring. SOL: sleep onset latency.

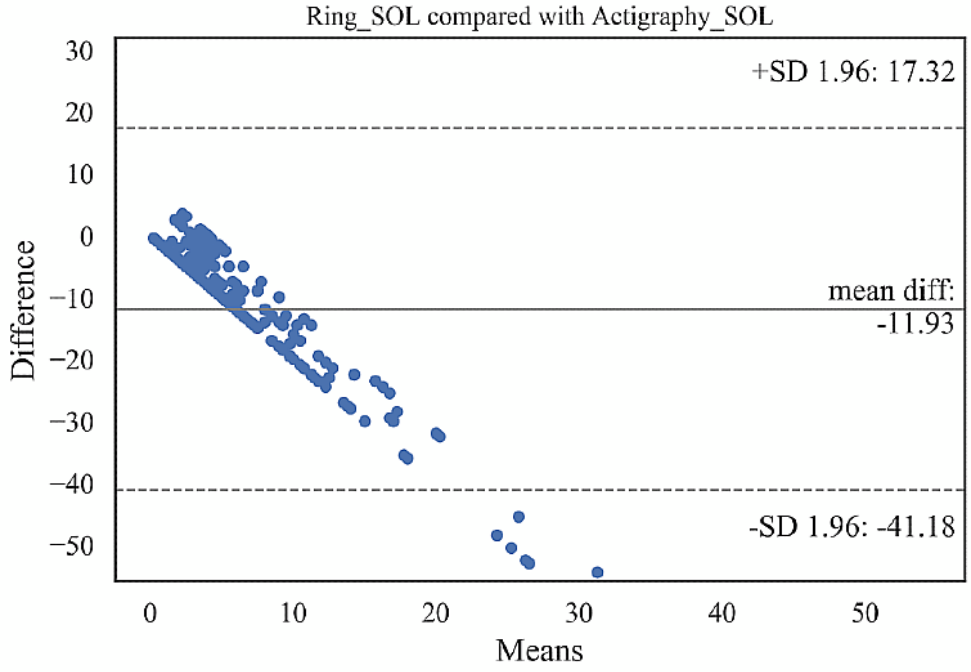


Figure 5. Bland-Altman plot for sleep onset latency estimated by the Samsung watch. SOL: sleep onset latency.

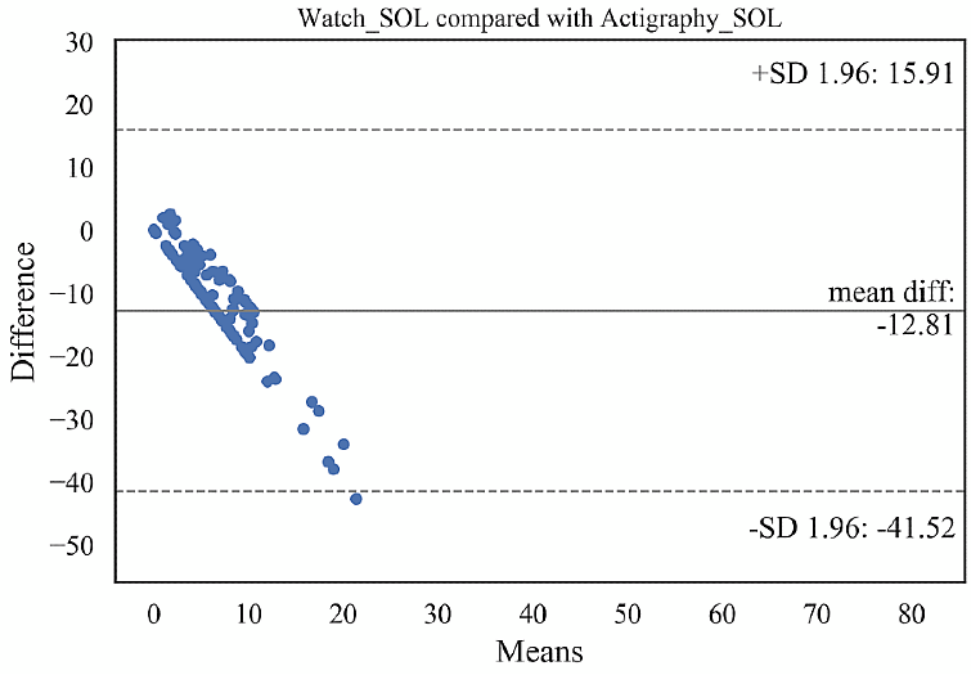


Table 11. Bias and agreement limits based on Bland-Altman plot of the sleep onset latency for the actigraphy and the Oura ring.

Parameter	Mean difference (SD)	Lower and upper agreement limits
Sleep onset latency (min)	−11.93 (14.92)	−41.18, 17.32

Table 12. Bias and agreement limits based on Bland-Altman plot of sleep onset latency for the actigraphy and Samsung watch.

Parameter	Mean difference (SD)	Lower and upper agreement limits
Sleep onset latency (min)	−12.81 (14.65)	−41.52, 15.91

The Oura ring overestimated the SOL, on average, by 11.93 min (95% C: −13.74 to −10.13) compared with the actigraphy. Out of 266 samples, 14 fell outside the agreement limits (lower limit −41.18 min, upper limit 17.32 min). Table 9 shows that the overestimation of the SOL by the ring was significant (t_{265}=−13.01; P<.001). Similarly, the watch overestimated the SOL, on average, by 12.81 min (95% CI −15.32 to −10.29). Most of the samples (all except 2) were within the agreement limits (lower limit −41.52 min, upper limit 15.91 min).

Discussion

Principal Findings

To the best of our knowledge, this is the first sleep validation study of the Oura ring and the Samsung watch performed under free-living conditions in comparison with an actigraphy method. The free-living condition allows participants to engage in their daily routines as usual during the monitoring. If commercial devices are used in trials under such free-living conditions, subjective evaluations and self-reports are insufficient to measure the validity of these devices [44-46]. It is important to test these devices against research devices to investigate their error margins and to standardize their software versions, minimizing controllable measurement differences. In contrast to related work, this study investigated wearables in a 1-week home-based monitoring, providing a higher confidence level on the validity of sleep parameters reported by these wearables. We discuss the results obtained and compare them with the related sleep validation studies, most of which are limited to the laboratory settings and compared with PSG.

Our findings showed that the mean differences of TST, WASO, and SE between the actigraphy device and the Oura ring were within the satisfactory range (ie, ≤30 min for TST and WASO and <5% for SE). Within the 266 valid total nights of sleep, only 14 TST, 17 WASO, and 18 SE fell outside the agreement limits. Our results also indicated significant correlations between the TST, WASO, and SE of the ring and the actigraphy. These findings are in accordance with a previous validation study of the Oura ring carried out in a single laboratory overnight study [35].

On the other hand, we found significant differences between the means of TST, WASO, and SE of the ring and the actigraphy. In our study, the Oura ring overestimated the TST (15.27 min) and underestimated the WASO (17.41 min) and SE (1.34%). Although the differences were within the satisfactory range, our results showed more overestimation and underestimation of the Oura ring than the lab-based sleep validation study [35]. This might be explained by the difference between the studies' samples and setups. Our study included more sleep data (ie, 225 more nights) and was performed in the house. Therefore, our results should be more accurate and have higher confidence levels in real-world applications. Unfortunately, these inaccuracies in sleep measurements in commercial devices might decrease their feasibility for clinical trials [47].

In accordance, the results showed biases in the sleep parameters provided by the Oura ring. However, the mean differences were within the satisfactory range, and only a few samples were outside the agreement limits. Therefore, the Oura ring can be acceptable for monitoring nonstaging sleep parameters under free-living conditions.

Moreover, our results indicated that the mean differences of the TST, WASO, and SE between the Samsung watch and the actigraphy were higher than the Oura ring's mean difference. The TST and SE mean differences of the watch were higher but still within the satisfactory range. However, the WASO mean difference (ie, 31.27 min) was negligibly higher than the range. Within the 134 valid total nights of sleep detection by the watch, 9 TST, 9 WASO, and 8 SE fell outside the agreement limit. Similarly, the correlation of the watch and actigraphy was lower than the ring, as the Pearson r values of the three parameters were closer to 0. Consequently, the sleep parameters of the watch had acceptable mean differences and indicated significant correlations with the actigraphy, but the Oura ring outperforms the Samsung watch in terms of the nonstaging sleep parameters.

Comparison With Prior Work

In previous studies, wrist activity trackers such as Fitbit Charge HR and Jawbone UP were compared with the PSG in lab tests on healthy adults [24,27,30]. The devices showed good agreement with the PSG in terms of TST, WASO, and SE. This is in accordance with our results for both the Oura ring and the Samsung watch. However, the overestimations or underestimations in our findings were higher than those in previous studies. The biases are particularly significant for the Samsung watch. For example, de Zambotti et al [24] indicated that the Fitbit Charge HR overestimates TST by 8 min and SE by 1.8% and underestimates WASO by 5.6 min. These low biases might be because of their limited setups and data collection, that is, an overnight laboratory sleep test on 32 healthy individuals.

There are a few studies performed under free-living conditions to evaluate activity trackers such as the Misfit Shine, Jawbone UP, and different models of Fitbit on healthy adults [23,48]. Our results regarding the Oura ring highlighted the high correlations obtained by these studies. For instance, Liang et al [48] indicated that there were high Pearson correlations between Fitbit Charge 2 and their baseline (a single-channel electroencephalogram-based device) in terms of TST (r=0.94), WASO (r=0.25), and SE (r=0.50). Ferguson et al [23] considered the TST correlations between four activity tracker devices and a research-grade accelerometer or multi-sensor device (BodyMedia SenseWear). The authors showed that the correlations were higher than 0.82 for the devices. On the other hand, our smartwatch results showed moderate correlations for TST, WASO, and SE.

Furthermore, we considered gender as a dependent variable to evaluate whether there was a mean difference in sleep parameter changes between male and female groups. Considering the Oura ring, our results showed a nonsignificant difference between female and male groups in TST, which is similar to the findings of de Zambotti et al [27]. Moreover, Carter et al [49] evaluated the objective estimation of sleep parameters compared with subjective assessments. In comparison with this study, we obtained similar results in terms of objective TST. However,

the watch in our study showed a significant difference in TST. Besides, both the ring and the watch indicated significant differences between female and male groups in WASO and SE, which disagrees with de Zambotti et al [27] but confirms the findings of Carter et al [49].

Limitations

We considered using an actigraphy device as the baseline method, which is one of the limitations of this study. Our analysis was limited to TST, WASO, and SE parameters. Although we collected the SOL of the Oura ring and the Samsung watch, we could not evaluate the values, as the SOL measure of the actigraphy is unreliable [14]. The actigraphy methods are insufficient for evaluation of sleep stages (eg, deep sleep). Therefore, future work should investigate the sleep stages provided by the ring and watch, considering a feasible PSG or electroencephalogram-based method designed for home-based monitoring.

Another limitation of this study is that only healthy participants were included in the analysis. However, other studies have shown that the accuracy of the wearables might differ for different population groups [29,34]. This issue may limit the generalizability of the findings. This study's future directions are to perform a home-based sleep validation study to assess the accuracy of wearables for population groups of different ages (eg, adolescents and older people) and sleep disorders (eg, obstructive sleep apnea). Besides, bed-based and ballistocardiograph-based sensors [50] can be used to mitigate user errors during data collection.

Conclusions

Sleep monitoring in free-living conditions becomes feasible and practicable using commercial devices such as smart rings and smartwatches. Notwithstanding the advances and feasibility of these wearables, their validity in terms of sleep parameters was not thoroughly investigated, especially for mid- to long-term studies in everyday settings. This study assessed the Oura ring and the Samsung Gear Sport watch by examining their TST, WASO, and SE under free-living conditions. The wearable devices were tested in home-based monitoring, where the sleep parameters of 45 healthy participants were tracked for 7 days. The assessment was performed in comparison with an actigraphy device, leveraging the paired *t* tests, Bland-Altman plots, and Pearson correlations. Sleep parameters were investigated considering the gender of the participants as a dependent variable. Our results showed that despite the statistically significant differences in the sleep parameters (ie, TST, WASO, and SE) of both the Oura ring and the Samsung watch compared with the actigraphy device, the mean differences were within the satisfactory ranges. The sleep parameters also indicated significant correlations with actigraphy. Besides, we showed that there was no significant difference in the validation of TST between male and female groups in the Oura ring; however, both the Oura ring and the Samsung watch indicated significant differences between the female and male groups in the estimation of WASO and SE.

Similarly, in a population sample of healthy adults, both the Oura ring and the Samsung watch had acceptable mean differences and indicated significant correlations with the actigraphy. However, the biases of the ring were considerably lower than the biases of the watch. Further validation is required to assess the validity of the sleep stages provided by the ring and the watch under free-living conditions. Moreover, future work should include the assessment of the devices for other population groups, such as individuals with sleep disorders.

Acknowledgments

The authors would like to thank Elisa Lankinen, Mohsen Saei Dehghan, Bushra Zafar, and Henrika Merenlehto for contributing to the data collection and Susanna Landström for contributing to the analysis of the ActiGraph data. This work was supported in part by the Academy of Finland through the Preterm Birth Prevention in Everyday Settings Project under the grants 313448 and 313449, in part by the Academy of Finland through the SLIM Project under the grants 316810 and 316811, and in part by the US National Science Foundation through the UNITE Project under the grant SCC CNS-1831918 and the WiFiUS grant CNS-1702950. All the data and diary templates used for supporting the results of this study may be obtained from the corresponding author upon request and signing of a data use agreement.

Conflicts of Interest

None declared.

References

1. 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]
2. Gupta R, Dahiya S, Bhatia M. Effect of depression on sleep: qualitative or quantitative? Indian J Psychiatry 2009 Apr;51(2):117-121 [FREE Full text] [doi: 10.4103/0019-5545.49451] [Medline: 19823630]
3. Besedovsky L, Lange T, Born J. Sleep and immune function. Pflugers Arch 2012 Jan;463(1):121-137 [FREE Full text] [doi: 10.1007/s00424-011-1044-0] [Medline: 22071480]
4. Tobaldini E, Fiorelli EM, Solbiati M, Costantino G, Nobili L, Montano N. Short sleep duration and cardiometabolic risk: from pathophysiology to clinical evidence. Nat Rev Cardiol 2019 Apr;16(4):213-224. [doi: 10.1038/s41569-018-0109-6] [Medline: 30410106]

5. Perez-Pozuelo I, Zhai B, Palotti J, Mall R, Aupetit M, Garcia-Gomez JM, 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]
6. Berry R, Wagner M. Sleep Medicine Pearls. Philadelphia, USA: Elsevier Health Sciences; 2014.
7. 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 1;26(3):342-392. [doi: 10.1093/sleep/26.3.342] [Medline: 12749557]
8. Roane BM, Van Reen E, Hart CN, Wing R, Carskadon MA. Estimating sleep from multisensory armband measurements: validity and reliability in teens. J Sleep Res 2015 Dec;24(6):714-721 [FREE Full text] [doi: 10.1111/jsr.12317] [Medline: 26126746]
9. Mantua J, Gravel N, Spencer RM. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. Sensors (Basel) 2016 May 5;16(5) [FREE Full text] [doi: 10.3390/s16050646] [Medline: 27164110]
10. 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]
11. Sánchez-Ortuño MM, Edinger JD, Means MK, Almirall D. Home is where sleep is: an ecological approach to test the validity of actigraphy for the assessment of insomnia. J Clin Sleep Med 2010 Feb 15;6(1):21-29 [FREE Full text] [Medline: 20191934]
12. Toon E, Davey MJ, Hollis SL, Nixon GM, Horne RS, Biggs SN. Comparison of commercial wrist-based and smartphone accelerometers, actigraphy, and PSG in a clinical cohort of children and adolescents. J Clin Sleep Med 2016 Mar;12(3):343-350 [FREE Full text] [doi: 10.5664/jcsm.5580] [Medline: 26446248]
13. Marino M, Li Y, Rueschman MN, Winkelman JW, Ellenbogen JM, Solet JM, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. Sleep 2013 Nov 1;36(11):1747-1755 [FREE Full text] [doi: 10.5665/sleep.3142] [Medline: 24179309]
14. Scott H, Lack L, Lovato N. A systematic review of the accuracy of sleep wearable devices for estimating sleep onset. Sleep Med Rev 2020 Feb;49:101227. [doi: 10.1016/j.smrv.2019.101227] [Medline: 31901524]
15. Kvedar J, Coye MJ, Everett W. Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff (Millwood) 2014 Feb;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992] [Medline: 24493760]
16. Qi J, Yang P, Min G, Amft O, Dong F, Xu L. Advanced internet of things for personalised healthcare systems: a survey. Pervasive and Mobile Computing 2017 Oct;41:132-149. [doi: 10.1016/j.pmcj.2017.06.018]
17. Azimi I, Rahmani AM, Liljeberg P, Tenhunen H. Internet of things for remote elderly monitoring: a study from user-centered perspective. J Ambient Intell Human Comput 2016 Jun 20;8(2):273-289. [doi: 10.1007/s12652-016-0387-y]
18. Gubbi J, Buyya R, Marusic S, Palaniswami M. Internet of Things (IoT): a vision, architectural elements, and future directions. Fut Generat Comput Syst 2013 Sep;29(7):1645-1660. [doi: 10.1016/j.future.2013.01.010]
19. Mieronkoski R, Azimi I, Rahmani AM, Aantaa R, Terävä V, Liljeberg P, et al. The Internet of Things for basic nursing care-a scoping review. Int J Nurs Stud 2017 Apr;69:78-90. [doi: 10.1016/j.ijnurstu.2017.01.009] [Medline: 28189116]
20. Koskimäki H, Kinnunen H, Kurppa T, Röning J. How Do We Sleep: a Case Study of Sleep Duration and Quality Using Data From Oura Ring. In: Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. 2018 Presented at: ACM'18; October 8-12, 2018; Singapore. [doi: 10.1145/3267305.3267697]
21. Lappalainen T, Virtanen L, Häkkilä J. Experiences With Wellness Ring and Bracelet Form Factor. In: Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia. 2016 Presented at: ACM'16; December 12-15, 2016; Rovaniemi, Finland. [doi: 10.1145/3012709.3016065]
22. Azimi I, Oti O, Labbaf S, Niela-Vilen H, Axelin A, Dutt N, et al. Personalized maternal sleep quality assessment: an objective iot-based longitudinal study. IEEE Access 2019;7:93433-93447. [doi: 10.1109/access.2019.2927781]
23. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 2015 Mar 27;12:42 [FREE Full text] [doi: 10.1186/s12966-015-0201-9] [Medline: 25890168]
24. de Zambotti M, Baker FC, Willoughby AR, Godino JG, Wing D, Patrick K, et al. Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. Physiol Behav 2016 May 1;158:143-149 [FREE Full text] [doi: 10.1016/j.physbeh.2016.03.006] [Medline: 26969518]
25. Moreno-Pino F, Porras-Segovia A, López-Esteban P, Artés A, Baca-García E. Validation of Fitbit Charge 2 and Fitbit Alta HR Against Polysomnography for Assessing Sleep in Adults With Obstructive Sleep Apnea. J Clin Sleep Med 2019 Nov 15;15(11):1645-1653. [doi: 10.5664/jcsm.8032] [Medline: 31739855]
26. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of wristband FitBit models in assessing sleep: systematic review and meta-analysis. J Med Internet Res 2019 Nov 28;21(11):e16273 [FREE Full text] [doi: 10.2196/16273] [Medline: 31778122]
27. de Zambotti M, Baker FC, Colrain IM. Validation of sleep-tracking technology compared with polysomnography in adolescents. Sleep 2015 Sep 1;38(9):1461-1468 [FREE Full text] [doi: 10.5665/sleep.4990] [Medline: 26158896]

28. de Zambotti M, Claudatos S, Inkelis S, Colrain IM, Baker FC. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chronobiol Int 2015;32(7):1024-1028 [FREE Full text] [doi: 10.3109/07420528.2015.1054395] [Medline: 26158542]
29. Cook JD, Prairie ML, Plante DT. Ability of the multisensory jawbone UP3 to quantify and classify sleep in patients with suspected central disorders of hypersomnolence: a comparison against polysomnography and actigraphy. J Clin Sleep Med 2018 May 15;14(5):841-848 [FREE Full text] [doi: 10.5664/jcsm.7120] [Medline: 29734975]
30. Spielmanns M, Bost D, Windisch W, Alter P, Greulich T, Nell C, et al. Measuring sleep quality and efficiency with an activity monitoring device in comparison to polysomnography. J Clin Med Res 2019 Dec;11(12):825-833 [FREE Full text] [doi: 10.14740/jocmr4026] [Medline: 31803327]
31. Meltzer LJ, Hiruma LS, Avis K, Montgomery-Downs H, Valentin J. Comparison of a commercial accelerometer with polysomnography and actigraphy in children and adolescents. Sleep 2015 Aug 1;38(8):1323-1330 [FREE Full text] [doi: 10.5665/sleep.4918] [Medline: 26118555]
32. Montgomery-Downs HE, Insana SP, Bond JA. Movement toward a novel activity monitoring device. Sleep Breath 2012 Sep;16(3):913-917. [doi: 10.1007/s11325-011-0585-y] [Medline: 21971963]
33. Kolla BP, Mansukhani S, Mansukhani MP. Consumer sleep tracking devices: a review of mechanisms, validity and utility. Expert Rev Med Devices 2016 May;13(5):497-506. [doi: 10.1586/17434440.2016.1171708] [Medline: 27043070]
34. Cook JD, Prairie ML, Plante DT. Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: a comparison against polysomnography and wrist-worn actigraphy. J Affect Disord 2017 Aug 1;217:299-305 [FREE Full text] [doi: 10.1016/j.jad.2017.04.030] [Medline: 28448949]
35. de Zambotti M, Rosas L, Colrain IM, Baker FC. The sleep of the ring: comparison of the ōura sleep tracker against polysomnography. Behav Sleep Med 2019;17(2):124-136 [FREE Full text] [doi: 10.1080/15402002.2017.1300587] [Medline: 28323455]
36. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
37. Oura Ring. URL: https://ouraring.com/ [accessed 2020-03-06]
38. Gear Sport. URL: https://www.samsung.com/us/explore/gear-sport/ [accessed 2020-03-05]
39. What does the 'Detect Sleep Periods' button do and how does it work? Actigraph. URL: https://actigraphcorp.force.com/support/s/article/What-does-the-Detect-Sleep-Periods-button-do-and-how-does-it-work [accessed 2020-10-18]
40. Cole RJ, Kripke DF, Gruen W, Mullaney DJ, Gillin JC. Automatic sleep/wake identification from wrist activity. Sleep 1992 Oct;15(5):461-469. [doi: 10.1093/sleep/15.5.461] [Medline: 1455130]
41. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
42. Watson PF, Petrie A. Method agreement analysis: a review of correct methodology. Theriogenology 2010 Jun;73(9):1167-1179 [FREE Full text] [doi: 10.1016/j.theriogenology.2010.01.003] [Medline: 20138353]
43. Meltzer L, Walsh C, Traylor J, Westin A. Direct comparison of two new actigraphs and polysomnography in children and adolescents. Sleep 2012 Jan 1;35(1):159-166 [FREE Full text] [doi: 10.5665/sleep.1608] [Medline: 22215930]
44. Girschik J, Fritschi L, Heyworth J, Waters F. Validation of self-reported sleep against actigraphy. J Epidemiol 2012;22(5):462-468 [FREE Full text] [doi: 10.2188/jea.je20120012] [Medline: 22850546]
45. Lauderdale DS, Knutson KL, Yan LL, Liu K, Rathouz PJ. Self-reported and measured sleep duration: how similar are they? Epidemiology 2008 Nov;19(6):838-845 [FREE Full text] [doi: 10.1097/EDE.0b013e318187a7b0] [Medline: 18854708]
46. 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]
47. Grym K, Niela-Vilén H, Ekholm E, Hamari L, Azimi I, Rahmani A, et al. Feasibility of smart wristbands for continuous monitoring during pregnancy and one month after birth. BMC Pregnancy Childbirth 2019 Jan 17;19(1):34 [FREE Full text] [doi: 10.1186/s12884-019-2187-9] [Medline: 30654747]
48. Liang Z, Chapa Martell MA. Validity of consumer activity wristbands and wearable EEG for measuring overall sleep parameters and sleep structure in free-living conditions. J Healthc Inform Res 2018 Apr 20;2(1-2):152-178. [doi: 10.1007/s41666-018-0013-1]
49. Carter JR, Gervais BM, Adomeit JL, Greenlund IM. Subjective and objective sleep differ in male and female collegiate athletes. Sleep Health 2020 Mar 5 epub ahead of print. [doi: 10.1016/j.sleh.2020.01.016] [Medline: 32147360]
50. Sadek I, Biswas J, Abdulrazak B. Ballistocardiogram signal processing: a review. Health Inf Sci Syst 2019 Dec;7(1):10. [doi: 10.1007/s13755-019-0071-7] [Medline: 31114676]

Abbreviations

PPG: photoplethysmogram
PSG: polysomnography
SE: sleep efficiency
SOL: sleep onset latency

TST: total sleep time
WASO: wake after sleep onset

Edited by G Eysenbach; submitted 19.05.20; peer-reviewed by Y Zhou, B Price, J Li, F Palmieri; comments to author 07.08.20; revised version received 04.09.20; accepted 23.09.20; published 02.11.20.

Please cite as:
Asgari Mehrabadi M, Azimi I, Sarhaddi F, Axelin A, Niela-Vilén H, Myllyntausta S, Stenholm S, Dutt N, Liljeberg P, Rahmani AM
Sleep Tracking of a Commercially Available Smart Ring and Smartwatch Against Medical-Grade Actigraphy in Everyday Settings: Instrument Validation Study
JMIR Mhealth Uhealth 2020;8(10):e20465
URL: http://mhealth.jmir.org/2020/10/e20465/
doi:10.2196/20465
PMID:33038869

Original Paper

Embodiment of Wearable Technology: Qualitative Longitudinal Study

Elizabeth C Nelson[1], MSc; Anneke M Sools[2], PhD; Miriam M R Vollenbroek-Hutten[1,3], PhD; Tibert Verhagen[4], PhD; Matthijs L Noordzij[2], PhD

[1]Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, Netherlands

[2]Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

[3]Ziekenhuis Groep Twente, Almelo, Netherlands

[4]Center for Market Insights, Amsterdam University of Applied Sciences, Amsterdam, Netherlands

Corresponding Author:
Elizabeth C Nelson, MSc
Faculty of Electrical Engineering, Mathematics and Computer Science
University of Twente
Drienerlolaan 5
7522 NB
Enschede
Netherlands
Phone: 31 642683161
Email: elizabeth@learnadaptbuild.com

Abstract

Background: Current technology innovations, such as wearables, have caused surprising reactions and feelings of deep connection to devices. Some researchers are calling mobile and wearable technologies cognitive prostheses, which are intrinsically connected to individuals as if they are part of the body, similar to a physical prosthesis. Additionally, while several studies have been performed on the phenomenology of receiving and wearing a physical prosthesis, it is unknown whether similar subjective experiences arise with technology.

Objective: In one of the first qualitative studies to track wearables in a longitudinal investigation, we explore whether a wearable can be embodied similar to a physical prosthesis. We hoped to gain insights and compare the phases of embodiment (ie, initial adjustment to the prosthesis) and the psychological responses (ie, accept the prosthesis as part of their body) between wearables and limb prostheses. This approach allowed us to find out whether this pattern was part of a cyclical (ie, period of different usage intensity) or asymptotic (ie, abandonment of the technology) pattern.

Methods: We adapted a limb prosthesis methodological framework to be applied to wearables and conducted semistructured interviews over a span of several months to assess if, how, and to what extent individuals come to embody wearables similar to prosthetic devices. Twelve individuals wore fitness trackers for 9 months, during which time interviews were conducted in the following three phases: after 3 months, after 6 months, and at the end of the study after 9 months. A deductive thematic analysis based on Murray's work was combined with an inductive approach in which new themes were discovered.

Results: Overall, the individuals experienced technology embodiment similar to limb embodiment in terms of adjustment, wearability, awareness, and body extension. Furthermore, we discovered two additional themes of engagement/reengagement and comparison to another device or person. Interestingly, many participants experienced a rarely reported phenomenon in longitudinal studies where the feedback from the device was counterintuitive to their own beliefs. This created a blurring of self-perception and a dilemma of "whom" to believe, the machine or one's self.

Conclusions: There are many similarities between the embodiment of a limb prosthesis and a wearable. The large overlap between limb and wearable embodiment would suggest that insights from physical prostheses can be applied to wearables and vice versa. This is especially interesting as we are seeing the traditionally "dumb" body prosthesis becoming smarter and thus a natural merging of technology and body. Future longitudinal studies could focus on the dilemma people might experience of whether to believe the information of the device over their own thoughts and feelings. These studies might take into account constructs, such as technology reliance, autonomy, and levels of self-awareness.

(JMIR Mhealth Uhealth 2020;8(11):e16973) doi:10.2196/16973

KEYWORDS

wearability; implantable wearable; body extension; smart prosthesis; implantable devices; technology dependence; cognitive prosthesis; phenomenology; embodied self-discrepancy; technology addiction; longitudinal qualitative design

Introduction

Individuals are increasingly wearing devices on their bodies, which monitor their behavior and provide coaching through associated apps. On one hand, the interaction with these types of devices might result in a sustained experience, where technology extends the body [1-3], cognition, and even self, which was recently conceptualized as wearable technology embodiment [4]. This is in line with other work in which wearables were referred to as cognitive prostheses [5,6]. With this label, the positive enabling effects of the wearable on cognition (eg, improved self-regulation for going to bed on time with sleep tracking) is put in the same category as those that can be expected in the motoric domain from a physical prosthesis. On the other hand, some researchers have suggested that the effects and even the active use of wearables are short lived (ie, not more than a few weeks) and inconsequential [7]. This might also be in line with a popular belief that wearables, such as an activity tracker that counts and displays steps, are devices that might have an initial appeal, but will quickly be discarded after the novelty effect wears off [7-11]. Importantly, it is unknown how people experience wearables over longer periods of time given the dearth of longitudinal qualitative research on this topic [8,12,13]. What is known from a qualitative and longitudinal perspective is how people engage with, adapt to, and embody physical prostheses [14-16]. Given the initial viewpoint presented above, the expectation for wearables is that there may be noteworthy overlap between user experiences with and perspectives on prosthesis and wearable embodiment. However, if wearables in practice are quickly discarded, the alignment with physical prosthesis embodiment (on which people rely heavily to regain physical capabilities) will be problematic or even impossible. The aim of this research was to study to what extent the experiences that previously have been associated with adaptation over several months to a physical prosthesis [14] are applicable and representative for the embodiment of a wearable.

Murray [14] investigated user perspectives on the embodiment of a physical prosthesis, which provided important insights into the phases of embodiment (ie, initial adjustment to the prosthesis) and the psychological responses to the prosthesis (ie, accepting the prosthesis as part of their body). The process phases and psychological responses represent a comprehensive foundation to understand and categorize embodiment of not only a prosthesis but also any item embodied by the user. Murray's six themes include (1) adjustment to a prosthesis; (2) balance of the body; (3) awareness of the prosthesis; (4) the prosthesis as a tool or corporeal structure; (5) the knowing body; and (6) the phantom becomes the prosthesis, extending the body.

The first theme (adjustment to a prosthesis) describes the initial period of mental and physical adjustment after receiving a new prosthetic device. In order to embody a prosthesis, an individual must create a working relationship with it, integrating it into the daily routine. The maintenance of the prosthesis during this time is described as considerable, but the tasks are eventually absorbed into a schedule or rhythm requiring little thought. Past research on technology use describes a similar process during the initial period with a new technology device. Research has found that interaction with a new technology first evokes a cognitive response to the device, followed by a behavioral response [17,18], which can either lead to the adoption or abandonment of the technology [19]. It is unknown to what extent initial technology adoption and experience is similar to physical prosthesis embodiment during the first 3 months of active use.

The second theme from Murray (the balance of the body) evaluates the adjustment to the imbalance created by the amputation or prosthetic device. Balance is key in creating a good fit between an individual and a prosthesis, making it easier to wear. Individuals wearing a well-balanced prosthetic device describe an ongoing process of "subconscious compensation" to naturally reposition the body to improve balance [14]. Creating a good fit with a prosthetic limb relates closely to the concept of wearability, which involves the degree of physical, mental, and/or social comfort in wearing a device [20]. Highly wearable technologies are comfortable and easy to wear (no distraction or attention demand) [20] and have been shown to have higher success rates for continued use [21]. Patients receiving a new wearable to replace an older version described feeling like "a living medical instrument" [22] while wearing the device. What is interesting to discover is to what extent the wearability of the wearable device impacts the embodiment of the device.

The third theme from Murray (awareness of the prosthesis) explains the changing nature of use over time (ie, disturbances or ease of use). A prosthesis that is embodied is integrated into the body, operating automatically without disruption or attention. Feedback from the technology, or contextual awareness [20], also has to feel automatic. Therefore, push notifications or other content can either be welcomed or be considered a disruption. The quality of the information and the extent to which people are engaged can ultimately impact their behavior [23,24]. Unpredictable or unusual feedback can engage people and compel their attention [20]. This presents a possible contrast between the unwelcomed interruption or awareness of a physical prosthesis and the welcome feedback interruption of a wearable.

Murray's fourth, fifth, and sixth themes analyze the deeper emotional relationship with the technology. Murray's fourth theme (the prosthesis as a tool or corporeal structure) examines whether individuals felt either a sense of completeness with the device or the prosthesis remained a helpful but external tool [14,25-28]. Research in mobile phones focuses on the adoption of technology beyond a technical tool, with features such as gamification [29,30]. Gamified elements have been successful

in short-term analysis [31-33] but have been critiqued by some for their questionable ability to aid in long-term goal attainment and behavior change [29]. Murray's fifth theme (the knowing body) involves the body's feedback to the mind. This can include the body's muscle memory where the feedback of an embodied prosthesis can be considered the same as the rest of the body. Murray's sixth theme (the phantom becomes the prosthesis, extending the body) explains where the prosthetic limb is considered as part of the body, possibly replacing the experience of a phantom limb (when the missing limb is felt). Interestingly, prosthesis research has shown that a phantom limb and a prosthesis can intertwine in an individual's mind, merging into one perceived entity [14]. All three themes examine the integral connection of the brain to the prosthesis. Research has suggested that technology may interlace with our minds to take over some tasks, such as navigation [34], and take on new ones, such as quantifying sleep and activity [4,26]. Clark [35] argued that cell phones were not simply technological tools but upgrades of the mind. Could the digital feedback intertwine with individuals' perceptions of their sleep or activity creating a combined feedback experience similar to the intertwining of a prosthesis and phantom limb? Additionally, what are the implications when it does?

While certain similarities have been established between wearable and physical prosthesis experiences, there are no longitudinal studies of wearables that can be compared with the prosthesis experience. Adapting the themes of Murray [14] to wearables can shed light on how a wearable may be embodied similar to a physical prosthesis in terms of both the phases of embodiment (ie, initial adjustment) and psychological responses (ie, accepting it as part of their body). This study extends current research to show usage patterns of a wearable over the long-term.

Methods

Participants

Over a 9-month period, a sample of 43 employees out of 400 from a large consultancy company in Amsterdam, The Netherlands, wore wearables, specifically the Jawbone UP Move, which consists of a wristband and mobile app, as well as a web platform with login (Figure 1) [36]. The wristband lit up and vibrated when a goal was met but did not have an interface or connection to any function or information beyond activity and sleep (ie, email, navigation, and phone calls). The mobile app included activity for the current day (including calories burned, idle time, distance traveled, and floors climbed) and time slept the previous night (including minutes asleep, minutes awake, and deep sleep and light sleep quality). The participants were part of a larger work wellness program and longitudinal research study. The wellness study was put together to experiment with improving health in the office and included aspects such as yoga classes, increased distribution of plants, and nutritious food. No incentives were provided, and there was no supervision of the group. People were not reminded about using the devices and could leave the study at any time.

Figure 1. Jawbone UP Move accelerometer and mobile phone app.

Of the 43 participants, 12 (mean age 35 years, SD 6.5 years, range 25-50 years; five female and seven male participants) volunteered to be part of this qualitative study. The wristbands were worn continuously by the participants and measured their sleep, activity, and inactivity. The mobile app was accessible on their smartphones and also over the internet with their personal login. Daily summaries were provided on the mobile app (Figure 1). Gamification elements in the app included duels (daily activity competitions) and daily goals of sleep and activity. Duels could be sent to other individuals in the workplace group to compete for the highest step count that day. The wristband vibrated and lit up when the daily activity goal was reached. Reminders were sent few (two to three) times per day regarding activity progress via the app. Data were transferred automatically from the wristband via Bluetooth to the mobile app and thus worked offline. Two of the male participants had worn a fitness tracker prior to the study. A signed consent form was collected at the beginning of the study, and verbal consent was obtained at the start of each interview.

Interviews

Participants were invited to the study via email. The series of semistructured interviews followed Murray's [14] method of interviewing three times over the course of 9 months. While the main questions remained the same over the three interviews, some probing was done to dive deeper into participants initial responses to questions. The interviews were conducted around 3 months, 6 months, and 9 months after the participants started wearing the wristbands. The final interview was conducted 6 to 8 weeks after the work wellness program concluded. A pretest

was conducted with four individuals from the study for comprehensiveness and clarity of the wording of the questions [37,38]. Minor changes were made to the wording of the questions based on feedback from the pretest for better comprehension or to stimulate a more thorough answer.

At the start of each interview, participants provided consent for the interview to be recorded. The interviewer explained to the participants that their responses would remain anonymous and would be used only for scientific publication. The participants were interviewed face to face with the exception of one phone call. Interviews were recorded and then transcribed into documents and coded according to participant (code name) and interview number (ie, interview 1, 2, or 3) [37]. Participants were Dutch natives and spoke English fluently. All interviews were conducted in English face to face with one member of the research team. No prior relationship between the interviewee and interviewer existed before the interviews began. Interviews were conducted in a quiet section of the office, and the duration ranged from 8 to 35 minutes depending on the participant's response length. Most participants responded openly and did not require much probing.

Question Adaptation

The prosthesis-based themes and questions from Murray's [14] research were adopted for use in this wearable technology study (Table 1 and Multimedia Appendix 1). Some themes were combined. The "adjustment to the prosthesis" and the "tool or corporal structure" themes were combined into one theme. The gamified tool was too intertwined with the initial process of adjusting to the prosthesis. The "knowing body" and the "phantom becomes the prosthesis" themes were also combined. While the differences between the knowing body and the phantom becomes the prosthesis themes are clear when applied to a physical prosthesis (ie, the mind and the body), the overlaps when applied to a wearable make it difficult to separate the two because both themes are in relation to the mind. An additional question was added to see how individuals compared using the wearable to their smartphone.

Table 1. Original themes of Murray and adapted themes with descriptions.

Original theme by Murray [14]	Description	Adapted theme	Description
Adjustment to the prothesis Tool or corporal structure	Becoming familiar with a prosthetic device for the first time, and physical and psychological adjustment. Experiencing the prosthetic device as either a tool or part of the body.	Adjustment to the wearable	Adjustment to the gamified tool/wearable during the initial period (months 1-3) with the device.
The balance of the body	Body weight distribution and balance.	Wearability	Level of comfort or ease of wear.
Awareness of the prosthetic device	The attention and awareness that was given to prosthesis use.	Awareness of the wearable	Level of awareness of the wearable and whether the aspects demanding awareness are welcome or disruptive.
The knowing body The phantom becomes the prosthesis, extending the body	The body's feedback to the mind (including the prosthetic body part). A prosthetic limb being experienced as part of the body.	The embodied wearable extending the mind	Experiencing the information as part of cognition and feeling and/or believing the information is as valid or more valid than subjective experience. Experiencing the device as part of the body.

Procedure and Analysis

In order to analyze the participant interviews, participants were given a pseudonym. Responses were then given the pseudonym as well as the interview sequence (in this case, interview 1, 2, or 3). ATLAS.ti (Scientific Software Development GmbH) and Microsoft Excel (Microsoft Corp) were used to analyze the responses for themes and longitudinal trends. Both inductive and deductive reasoning were used. Responses were coded into one or more of the adapted themes of Murray [14]. To stay open to additional themes, a constant comparison method [39,40], a form of thematic analysis [41], was utilized throughout the series of interviews to understand how participants experienced the wearable. This particular form of qualitative analysis was selected because its approach is useful in shedding light on how individuals experience the technology over the long term [14,38]. This dual process of constant comparison allowed us to code responses based on predefined categories, as well as discover themes or categories from the data itself.

Results

Themes

During analysis, two additional themes were discovered and included. A high number of responses were recognized relating to engagement/reengagement and comparison to another person or device. No additional questions were added to the semistructured interviews. While past studies have compared new prostheses to old prostheses, it was not a theme of Murray [14]. We believed these discovered themes do shed light on the nuances of wearable technology embodiment not yet discovered. The interrater reliability was calculated for all the themes using Cohen κ (Multimedia Appendix 2), with a 10% sample of the responses. Two researchers (ECN and MLN) reached a Pr(a) of 0.84 for all themes (Multimedia Appendix 2) after three rounds of revisions and thus a strong level of consistency [42]. It was established that the data collection did reach saturation.

Theme 1: Adjustment to the Wearable

During the initial 2-month period with the wearable, the participants described near constant interaction with the technology, checking it multiple times per day and sometimes only a few minutes apart to see how many steps had been achieved. Participants reported that the first months were "very motivating" and could be attributed to being drawn into the gamified elements of the technology. Most of the participants were experiencing a wearable for the first time (10 out of 12). Many individuals reported that they started walking to other parts of the office to have "face-to-face check-ins instead of sending a text or email." The technology also awarded digital badges to participants who challenged other participants to a "duel" and logged the most steps that day. However, these challenges or duels were reported to greatly decrease over time. All participants reported changes in their behavior and in making decisions to try to get more steps and/or to sleep longer or more deeply because of such aspects as the daily goal. One participant responded as follows:

> *I think in the beginning it's more of a high that you really want to achieve 10,000 steps.* [Eva; first interview]

While most of the responses were positive, there were some negative reactions to the recommended daily goals. The app suggested 10,000 steps and 8 hours of sleep. The suggested activity goal was explained as an "exciting challenge" in the beginning but proved to be "quite difficult" during weekdays, causing frustration and sometimes demotivation. Forgetting the wearable was also reported as frustrating especially during active days that could have increased the weekly average. In the first interview, Nate stated that he took 3000 steps on a normal weekday at the office. He expressed frustration with having to do more than three times that number to reach his daily goal. He made the following statement:

> *That's really crazy then it's difficult to get your 10k steps which I did as my target but it's really difficult to get there and I don't think you can get there with a normal job.* [Nate; second interview]

This initial adjustment time was also when the participants stated they "set up a routine" with the device, including pressing the button after waking in the morning and before sleep at night, learning the number of steps in typical activities, such as walking to work and usual errands, and checking step count throughout the day. Murray described this period as a time of acceptance or rejection when the individual and the prosthetic device must synchronize to achieve a working partnership. The participants reported a range of feelings to the new device such as feeling "familiarity," "curiosity," and "adjustment." These feelings show a similar pattern to prosthesis embodiment and suggest that the devices are embodied or are in the process of embodiment.

Theme 2: Engagement/Reengagement

While the first period of adjustment was also the period of most frequent use for all participants (n=12), overall, we saw a process of engaging, disengaging, and reengaging over the long term. Over the 9 months, most participants experienced at least two distinct periods of heavy use and two periods of infrequent use. One participant commented as follows:

> *I was very curious. How does it work and how much do I walk and now it became just a part of the day.* [Eliz; first interview]

Other participants described "missing the technology" when not wearing it, and a sense of "starting over" when the technology was not used for a period of time. Matthew described a period of infrequent use and had experienced sleep deprivation and reduced activity. This period was during a time of intense workload and long hours. He described the challenge of reintegrating the device into his daily routine again, where increased use helped him to regulate his routine once again. His comment was as follows:

> *Using the device again it feels like some kind of start over.* [Matthew; second interview]

One male participant in the first interview reported being especially interested in the sleep patterns and the quality of sleep, putting much focus on it. It was something that he had "not focused on" before getting the device. By the third interview, he believed he had learned to measure his sleep independent of the device. Interestingly, when participants were asked if they preferred the technology quantifying their activity and sleep or desired to gain the skill of knowing their quantified health data, most (n=9) preferred to continue using the technology. The reengagement by all participants and desire to continue using the technology as opposed to gaining the skill suggests a certain level of embodiment of the device.

Theme 3: Wearability

The wearable in this study was quite small compared with others on the market, so neither women nor men complained about its bulkiness. In general, participants in this study described the device as "comfortable." The participants received the device in the autumn and initially talked about adjusting to sleeping with it, but most (n=10) said they "did not notice" the wearable or that it "did not bother them." One participant chose not to sleep with the wearable on finding it uncomfortable after a few nights. When the spring and summer months came, some participants (n=2) reported that the device became "itchy" during high heat. A participant who also struggled initially with discomfort at night reported added discomfort during the summer in high heat. Interestingly, many participants described the experience similar to wearing a watch and stated that the device was "hardly noticeable." Yet, most participants did not wear a watch and had not used a watch for many years. No participant stopped wearing the wearable completely owing to comfort issues. One participant made the following statement:

> *It felt like I was wearing a watch. In winter, it was ok but in summer I thought it was sometimes a bit annoying.* [Mary; third interview]

The level of perceived attractiveness seemed to add to the ease of wear. Most participants (n=10) found the device attractive, and many (n=9) also enjoyed being asked what the device was. This positive attention and perceived attractiveness made the device quite wearable. Interestingly, by the end of the study, many participants stated a desire for their next wearable to have

"more functionality" and a different look (n=9), and to be more like a smartwatch than an activity tracker. One participant made the following statement:

> *I don't really feel it. It doesn't bother me at all. My only thing is that it's ugly.* [Mathew; first interview]

Wearability thus seemed to focus heavily on seasonality, the size and feeling of a watch, discomfort during sleep or high heat, and the level of perceived attractiveness.

Theme 4: Awareness of the Wearable

After the initial 2 months, many participants interacted much less with the device. The daily results and content from the mobile app (Figure 1) were explained as "repetitive" and "did not change or surprise" the participants anymore. John initially reported being "addicted to the device" and did not experience the first low interaction period until much later than the other participants. However, by the third interview, he described a feeling of "boredom with the content" of the mobile app. The majority of the participants did not have a high level of awareness of the device stating it was "just there." John made the following statements:

> *You were constantly looking how many steps I've taken in the last 10 minutes and now it's become like I said hardly think of that it's there… you're being reminded because people keep asking what's that on your wrist. But not because I feel it, sense it. It's just there.* [John; first interview]

> *There are also some notification or suggestions for what you can do you just can't help reading them so when they pop up you see them and obviously see there is a pattern or there are certain standard suggestions once you've seen them.* [John; second interview]

All participants reported checking sleep and activity once in the morning for sleep and at the end of the day for activity after the initial period of high use. The wearables required a button to be pushed before bed and when awake in the morning as an extra framework to accurately measure sleep time. The wearable did not demand attention similar to a well-fit prosthesis, but the feedback was considered boring (ie, low contextual awareness).

Theme 5: The Embodied Wearable, Extending the Mind

Many participants expressed "intense" reactions to or relationships with their device during the 9-month study and described "missing the device" when it was forgotten. While some reported experiences of "addiction," at minimum, reports referred to the technology as likely to be "habit forming." Four participants said the device felt like part of their body. Some participants made the following statements:

> *There won't be many moments that I forget about it or won't wear it. I take it to the gym it's really part of my body.* [Anna; second interview]

> *I'm quite surprised that it's become such an automatic, almost part of your body so to say. Like I said I don't feel it, I don't notice it, so it's there.* [John; first interview]

In the second interview, John described what he may do moving forward. He might either stop using the device or keep it for his physical training to "keep himself sharp." By the third interview, the wearable was replaced with another focused on running training, which was a gift from his team. He enjoyed the additional information like heart rate. In Tom's first interview, he described feeling powerless and wanting to know his biological data when he was without the device. He stated "It's a crazy feeling when it's off." He restated this in the second interview and mentioned "because you get comfortable wearing it, so when you take it off you miss it." In the third interview, he seemed to refer to himself and the technology working together.

Many participants referenced the calming effects confirming health behavior but could not answer why. One participant stated in all three interviews that it was comforting knowing what the activity had been. This was true for reaching milestones, such as 10,000 steps, and when the device confirmed feeling tired, such as after a bad night's sleep. The adverse reaction existed when goals were not met and feelings of discomfort or frustrations arose. One participant commented as follows:

> *Your feeling is being confirmed. If you think it was a rough night then you look at the app "oh it was".* [Eliz; second interview]

All participants experienced "surprising" and "discomforting" feedback regarding their sleep and/or activity. Some participants (n=4) reported a shift later in the study and started questioning the technology as an "accurate/correct measurement" while continuing to use it. They seemed to be struggling to decide whether to believe themselves or the device. Interestingly, none of the four participants reached a final decision on "whom" to trust by the end of the study. The indecision and willingness to trust the wearable above one's own feeling would suggest some embodiment had taken place.

Theme 6: Comparison to Another Device or Person

A new theme was discovered based on comparisons to other devices. This theme was further explored using responses from the adapted questions, such as how the wearable compared with the smartphone. Half of the participants (n=6) found the experience similar to receiving their first smartphone and checking their wearable "automatically" and with "little thought," although this statement was not consistent for individuals over the three interviews (Multimedia Appendix 3). This was explained by the fact that the wearable, like the phone, was "constantly with them," gave "updated data," and was something they "checked often." The experience was also compared to that of a singular mobile app on a phone but was felt to be less relevant than the full functionality of a smartphone and was therefore less frequently used. This was believed to be a reaction to the evolution of technology and the initial experience of exploration with a new device. One participant made the following statement:

> *It's not like my phone. It's a pull it's not addictive and disappointment is maybe not the correct word. It's more like you want to wear it because you have it and*

> *it doesn't take you a lot of effort to have it record all the steps.* [Tom; third interview]

One group of participants (n=8) had a reliance on the technology for activity and sleep evaluation, which they explained as similar to a smartphone. While most participants found they could not know their sleep and activity without the device, they reported at least some level of learning to gauge activity and sleep. The wearable took the place of a guide, even being referenced to as a "mother" because of the reminders to be healthier. One participant made the following statement:

> *Every morning you get the alert about the notification of your sleep. It's good to see and I track how much is my sleep and is it long enough sound sleep. I want to try more sound sleep than light sleep because I sleep all night but more lite than sound sleep. But it's funny, I call the notification to go to sleep 'mother'.* [Matthew; second interview]

Furthermore, most of our participants stated that they depended on their phone for navigation (n=11) and phone numbers (n=12) [43]. In the first interview, one participant stated that there was "no way of realistically guessing progress." By the third interview, the participant stated that it may be possible to guess progress after using the device for so many months, but still felt the device was more accurate and motivating.

Discussion

Principal Findings

This study is one of the first to track the use of wearables in a longitudinal qualitative study, providing a nuanced and varied insight into how this technology is used, embodied, and integrated into people's daily life. In this study, we found that although previous short-term research seemed to suggest wearables are quickly abandoned [7,8,10,11], there were alternating periods of engagement, disengagement, and reengagement over a 9-month timeframe. The fluctuating engagement is consistent with Oliver's [44] belief that the usage process of a device includes a sequence of acceptance, experience, verification, and continued use. On further reflection, it is understandable that engagement and reengagement did not come up in research on prosthetic limbs. Limb prostheses are associated with an experience of acceptance of the prosthetic device but rarely periods of reengagement, because this would mean giving up the regained functions. Furthermore, a limb would not likely be compared to another person.

It is also important to note that unlike the report by Murray [14], our participants were not obliged to wear the wearable, yet none of the participants chose to leave the study. Many participants expressed desires for higher functioning devices as their next wearable, suggesting a continued interest in the device/process. The most frequent use was experienced by all participants at the beginning of the study. This was during the period where many described getting the device into their daily routine and enjoying the gamified elements of the device, including competition with friends or colleagues. This high level of motivation is experienced often with technology adoption [17], especially technology including gamification [21], and is explained by theoretical paradigms, such as the innovation diffusion theory, which describes how beliefs, such as relative advantage, influence how individuals decide to adopt a technology [45,46].

We also discovered differences and similarities between a wearable and a limb prosthesis. With both, successful adjustment to the device during the first period and wearability are key to adoption and embodiment. High wearability means the device should be comfortable and should integrate into the body to fit within the individual's overall functioning without constraining any motion [47]. The wearable we chose was relatively small, and if it had been larger or heavier, or had hindered movement, the embodiment could have been less successful [48]. Itchiness in warmer seasons did seem to provide some discomfort. Wearability is especially important for wearables that are an optional addition to the body, while a limb prosthesis is considered a needed replacement. The participants' positive reactions to public approval are supported by the Theory of Planned Behavior [49]. Initial reactions to the devices also suggest some form of empowerment or perceived behavioral control [21]. Subjective norms can pressure individuals for approval, changing their behavior or perceptions. Wearing a device that received positive attention could reinforce the choice to wear the device [48] and perceive the device as more wearable due to the emotional benefits. Wearables also differed from Murray's [14] prosthesis research in terms of awareness. The wearable did not create bothersome awareness similar to a well-fit prosthesis but was not considered engaging in the long term (ie, low contextual awareness).

The feedback from the wearable created an interesting dilemma. Adding a quantified measurement of sleep and activity to the perceptions of sleep and activity created parallel feedback that could be either confirming (ie, confirm a good or bad night's sleep) or invalidating (ie, present information radically different than experiential perceptions). Parallel feedback can create a fracture in the sense of self and can lead to either distrusting the device, one's self, or both [50]. When the individual trusts the information coming from the device, this can manifest in self-regulating responses. However, when the information does not match the individual's beliefs, this can cause a reaction of self-discrepancy, where the individual holds simultaneous yet incompatible self-beliefs possibly causing distress [51]. A merging experience has been described in technology embodiment research [2,3] and cyborg intentionality where the person and technology both come with separate intentions (mediated intentionality/composite intentionality) combined into an experience [1]. For instance, the wearable did not intend to be a "mother figure," yet the combination of individuals' experiences with the wearable's intention created that "human dynamic." We did not find that participants experienced the discrepancy as particularly distressing, but awareness of the discrepancy seemed to increase over time. Interestingly, individuals more boldly expressed the inconsistency as time went on, but these individuals continued to struggle with whether to abandon the device. There were clear individual differences in the extent to which people experienced discrepancies and the willingness to put more trust in the

technologically mediated information than in subjective experience. The experience of seeing the device as a "mother figure" may show both negative and positive reactions to self-regulation. Further longitudinal investigation of this phenomenon has been called for to see how real-time monitoring of human functions and constant presentation of the data influence self-perception [50]. The self-discrepancy theory has not yet been extended to embodiment literature. We believe this research provides the first evidence that the self-discrepancy theory is applicable to embodiment phenomena.

In addition to the adapted themes of Murray [14], we discovered that wearables do have similarities to other devices such as mobile phones. We see increasing dependence on technologies, such as mobile phones and navigation technologies, that extend or replace our cognitive efforts and competencies (ie, remembered phone numbers and navigation) [6,34,35]. Research in mobile addiction or "smartphone dependence" has addressed its effect on mental health issues with regard to compulsive usage [52]. Our comparison of wearables to initial smartphone experiences suggests that wearables could have similar reactions to the technology, especially as the technology innovates. Research has found that individuals can create an emotional attachment to their mobile phones [53], and literature within mobile technology interaction has proven that individuals can believe their phones are extensions of themselves [4,8,54,55].

Limitations and Future Research

This study has few limitations. First, the study included a relatively small and homogenous group (n=12) that was repeatedly measured over a period of 9 months, making further generalization a question for future research. The participant group was representative of a consumer group (young to middle-aged highly educated professionals) interested in wearables. For many of the participants, this was their first time using wearables, which provides great insights into the experiences of first-time users. However, we recommend performing a further study on participants having ongoing experience with wearables and mobile devices. The wearables were small and unobtrusive, and while this helped us to see the experience of a highly wearable device, we acknowledge that not all wearables are unobtrusive. Additionally, while all participants could abandon the study at any point, they were part of a wellness program at work, which may have encouraged them to continue. Our discovered theme of comparison to another device or person could indicate levels of technology dependence similar to mobile phones. We recommend further research on technology dependence and addiction to various types of technologies. Furthermore, this study did not examine or provide explicit information to participants about the validity and reliability of the consumer wearables. This is an important research topic in and of itself [56], and for this study, the participants were referred to the disclaimer of the company indicating the device was not a medical grade device and was only intended for lifestyle monitoring and suggestions. We also acknowledge the critique on phenomenological studies more generally [57] that language mediates experience, and hence, a direct window on the experience itself was not our aim. Finally, this study compared a typical "dumb" physical prosthesis to a "smart" wearable, which are two areas already merging into a shared space. This research provides important insights into the experience of future "smart" prostheses and technologies already embodied by our bodies and minds [4].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview questions.
[DOCX File , 15 KB - mhealth_v8i11e16973_app1.docx]

Multimedia Appendix 2
Cohen kappa.
[DOCX File , 17 KB - mhealth_v8i11e16973_app2.docx]

Multimedia Appendix 3
Comparisons.
[DOCX File , 23 KB - mhealth_v8i11e16973_app3.docx]

References

1. Verbeek P. Cyborg intentionality: Rethinking the phenomenology of human–technology relations. Phenom Cogn Sci 2008 Jun 26;7(3):387-395. [doi: 10.1007/s11097-008-9099-x]
2. Ihde D. Technology and the Lifeworld: From Garden to Earth. Bloomington, IN: Indiana University Press; 1990.
3. Ihde D. Existential Technics. Albany, NY: State University of New York Press; 1985.
4. Nelson EC, Verhagen T, Vollenbroek-Hutten MM, Noordzij ML. Is Wearable Technology Becoming Part of Us? Developing and Validating a Measurement Scale for Wearable Technology Embodiment. JMIR Mhealth Uhealth 2019 Aug 09;7(8):e12771 [FREE Full text] [doi: 10.2196/12771] [Medline: 31400106]

5. De Preester H. Technology and the Body: the (Im)Possibilities of Re-embodiment. Found Sci 2010 Sep 26;16(2-3):119-137. [doi: 10.1007/s10699-010-9188-5]
6. Sarosiek A. New technologies as prosthesis of cognitive system. Semina Scientiarum 2016;15:64-76. [doi: 10.15633/ss.1768]
7. Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, et al. Effect of Wearable Technology Combined With a Lifestyle Intervention on Long-term Weight Loss: The IDEA Randomized Clinical Trial. JAMA 2016 Sep 20;316(11):1161-1171 [FREE Full text] [doi: 10.1001/jama.2016.12858] [Medline: 27654602]
8. Sultan N. Reflective thoughts on the potential and challenges of wearable technology for healthcare provision and medical education. International Journal of Information Management 2015 Oct;35(5):521-526. [doi: 10.1016/j.ijinfomgt.2015.04.010]
9. Ledger D, McCaffrey D. Inside Wearables Part 1: How behavior change unlocks long-term engagement. Medium. 2014. URL: https://medium.com/@endeavourprtnrs/inside-wearable-how-the-science-of-human-behavior-change-offers-the-secret-to-long-term-engagement-a15b3c7d4cf3 [accessed 2018-04-16]
10. Emerging mHealth: Paths for growth. PwC. 2014. URL: https://www.pwc.com/gx/en/healthcare/mhealth/assets/pwc-emerging-mhealth-full.pdf [accessed 2020-10-22]
11. Canhoto AI, Arp S. Exploring the factors that support adoption and sustained use of health and fitness wearables. Journal of Marketing Management 2016 Oct 26;33(1-2):32-60. [doi: 10.1080/0267257x.2016.1234505]
12. Spruijt-Metz D, Nilsen W. Dynamic Models of Behavior for Just-in-Time Adaptive Interventions. IEEE Pervasive Comput 2014 Jul;13(3):13-17. [doi: 10.1109/mprv.2014.46]
13. Nelson EC, de Keijzer R, Vollenbroek-Hutten MM, Verhagen T, Noordzij ML. The Relationship Between Diagnosed Burnout and Sleep Measured by Activity Trackers: Four Longitudinal Case Studies. In: Body Area Networks: Smart IoT and Big Data for Intelligent Health Management. 2019 Presented at: Bodynets; October 2-3, 2019; Florence, Italy. [doi: 10.1007/978-3-030-34833-5_24]
14. Murray CD. An interpretative phenomenological analysis of the embodiment of artificial limbs. Disabil Rehabil 2004 Aug 19;26(16):963-973. [doi: 10.1080/09638280410001696764] [Medline: 15371044]
15. Merleau-Ponty M. Phenomenology of Perception. Abingdon, United Kingdom: Routledge; 2002.
16. Raspopovic S, Capogrosso M, Petrini FM, Bonizzato M, Rigosa J, Di Pino G, et al. Restoring natural sensory feedback in real-time bidirectional hand prostheses. Sci Transl Med 2014 Feb 05;6(222):222ra19. [doi: 10.1126/scitranslmed.3006820] [Medline: 24500407]
17. Brown SA, Dennis AR, Venkatesh V. Predicting Collaboration Technology Use: Integrating Technology Adoption and Collaboration Research. Journal of Management Information Systems 2014 Dec 08;27(2):9-54. [doi: 10.2753/mis0742-1222270201]
18. Goodhue DL, Thompson RL. Task-Technology Fit and Individual Performance. MIS Quarterly 1995 Jun;19(2):213-288. [doi: 10.2307/249689]
19. 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]
20. Dunne L, Smyth B. Psychophysical Elements of Wearability. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2007 Apr Presented at: SIGCHI Conference on Human Factors in Computing Systems; 2007; San Jose, CA p. 299-302. [doi: 10.1145/1240624.1240674]
21. Nelson EC, Verhagen T, Noordzij ML. Health empowerment through activity trackers: An empirical smart wristband study. Computers in Human Behavior 2016 Sep 30;62:364-374. [doi: 10.1016/j.chb.2016.03.065]
22. Fensli R, Boisen E. Human Factors Affecting the Patient's Acceptance of Wireless Biomedical Sensors. In: Biomedical Engineering Systems and Technologies. 2008 Presented at: International Joint Conference on Biomedical Engineering Systems and Technologies; January 2008; Berlin, Heidelberg p. 402-412. [doi: 10.1007/978-3-540-92219-3_30]
23. Appelboom G, Yang AH, Christophe BR, Bruce EM, Slomian J, Bruyère O, et al. The promise of wearable activity sensors to define patient recovery. J Clin Neurosci 2014 Jul;21(7):1089-1093. [doi: 10.1016/j.jocn.2013.12.003] [Medline: 24534628]
24. Chittaro L, Sioni R. Affective computing vs. affective placebo: Study of a biofeedback-controlled game for relaxation training. International Journal of Human-Computer Studies 2014 Aug;72(8-9):663-673. [doi: 10.1016/j.ijhcs.2014.01.007]
25. Scarry E. The merging of bodies and artifacts in the social contract. Semantic Scholar 1994;15.
26. De Preester H, Tsakiris M. Body-extension versus body-incorporation: Is there a need for a body-model? Phenom Cogn Sci 2009 Feb 27;8(3):307-319. [doi: 10.1007/s11097-009-9121-y]
27. Giummarra MJ, Gibson SJ, Georgiou-Karistianis N, Bradshaw JL. Mechanisms underlying embodiment, disembodiment and loss of embodiment. Neurosci Biobehav Rev 2008;32(1):143-160. [doi: 10.1016/j.neubiorev.2007.07.001] [Medline: 17707508]
28. Martel M, Cardinali L, Roy AC, Farnè A. Tool-use: An open window into body representation and its plasticity. Cogn Neuropsychol 2016;33(1-2):82-101 [FREE Full text] [doi: 10.1080/02643294.2016.1167678] [Medline: 27315277]
29. Tang L, Kay J. Gamification: metacognitive scaffolding towards long term goals? 2014 Presented at: UMAP Workshops; July 7-11, 2014; Aalborg, Denmark p. 63.
30. Mirvis P, Csikszentmihalyi M. Flow: The Psychology of Optimal Experience. The Academy of Management Review 1991 Jul;16(3):636-640. [doi: 10.2307/258925]

31. Foster S, Esper S, Griswold W. From Competition to Metacognition: Designing Diverse, Sustainable Educational Games. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 2013; Paris, France p. 99-108. [doi: 10.1145/2470654.2470669]
32. Goehle G. Gamification and Web-based Homework. PRIMUS 2013 Mar;23(3):234-246. [doi: 10.1080/10511970.2012.736451]
33. Sabourin J, Shores L, Mott B, Lester J. Predicting student self-regulation strategies in game-based learning environment. In: Intelligent Tutoring Systems. 2012 Presented at: International Conference on Intelligent Tutoring Systems; June; Berlin, Heidelberg p. 141-150. [doi: 10.1007/978-3-642-30950-2_19]
34. Ishikawa T, Fujiwara H, Imai O, Okabe A. Wayfinding with a GPS-based mobile navigation system: A comparison with maps and direct experience. Journal of Environmental Psychology 2008 Mar;28(1):74-82. [doi: 10.1016/j.jenvp.2007.09.002]
35. Clark A. Natural-Born Cyborgs? In: Beynon M, Nehaniv CL, Dautenhahn K, editors. Cognitive Technology: Instruments of Mind. CT 2001. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer; 2001:17-24.
36. Broughall N. Jawbone UP Move review. Techradar. 2014. URL: https://www.techradar.com/reviews/gadgets/jawbone-up-move-1277383/review [accessed 2019-03-06]
37. Brinkmann S. Methodological breaching experiments: Steps toward theorizing the qualitative interview. 2016 Jul 24;22(4):520-533. [doi: 10.1177/1354067X16650816]
38. Smith JA, Jarman M, Osborn M. Doing Interpretative Phenomenological Analysis. In: Qualitative Health Psychology: Theories and Methods. Thousand Oaks, CA: SAGE Publications; 1999:218-240.
39. Creswell JD, Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA: SAGE Publications; 2003.
40. Okazaki S, Mendez F. Perceived Ubiquity in Mobile Services. Journal of Interactive Marketing 2013 May;27(2):98-111. [doi: 10.1016/j.intmar.2012.10.001]
41. Giorgi A. The Descriptive Phenomenological Psychological Method. J Phenomenol Psychol 2012;43(1):3-12. [doi: 10.1163/156916212x632934]
42. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012;22(3):276-282. [doi: 10.11613/bm.2012.031]
43. Adawi M, Bragazzi NL, Argumosa-Villar L, Boada-Grau J, Vigil-Colet A, Yildirim C, et al. Translation and Validation of the Nomophobia Questionnaire in the Italian Language: Exploratory Factor Analysis. JMIR Mhealth Uhealth 2018 Jan 22;6(1):e24 [FREE Full text] [doi: 10.2196/mhealth.9186] [Medline: 29358161]
44. Oliver RL. A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. Journal of Marketing Research 2018 Nov 28;17(4):460-469. [doi: 10.1177/002224378001700405]
45. Agarwal R. Individual acceptance of information technologies. In: Framing the Domains of IT Management: Projecting the Future...Through the Past. Cincinnati, OH: Pinnaflex Educational Resources; 2000:85-104.
46. Lee YH, Hsieh YC, Hsu CN. Adding Innovation Diffusion Theory to the Technology Acceptance Model: Supporting Employees' Intentions to use E-Learning Systems. 2011;14(4):124-137 [FREE Full text]
47. Prattichizzo D, Chinello F, Pacchierotti C, Malvezzi M. Towards wearability in fingertip haptics: a 3-DoF wearable device for cutaneous force feedback. IEEE Trans Haptics 2013;6(4):506-516. [doi: 10.1109/TOH.2013.53] [Medline: 24808402]
48. van Lier HG, Oberhagemann M, Stroes JD, Enewoldsen NM, Pieterse ME, Schraagen JM, et al. Design Decisions for a Real Time, Alcohol Craving Study Using Physio- and Psychological Measures. In: Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors. 2017 Presented at: International Conference on Persuasive Technology; April 4-6, 2017; Amsterdam, The Netherlands p. 3-15. [doi: 10.1007/978-3-319-55134-0_1]
49. Ajzen I. The theory of planned behaviour is alive and well, and not ready to retire: a commentary on Sniehotta, Presseau, and Araújo-Soares. Health Psychol Rev 2015;9(2):131-137. [doi: 10.1080/17437199.2014.883474] [Medline: 26209198]
50. Fairclough SH. Fundamentals of physiological computing. Interacting with Computers 2009 Jan;21(1-2):133-145. [doi: 10.1016/j.intcom.2008.10.011]
51. Higgins ET. Self-discrepancy: A theory relating self and affect. Psychological Review 1987;94(3):319-340. [doi: 10.1037/0033-295x.94.3.319]
52. Koufaris M. Applying the technology acceptance model and flow theory to online consumer behavior. Information Systems Research 2002(1):205-223 [FREE Full text]
53. Vincent J. Emotional attachment and mobile phones. Know Techn Pol 2006 Mar;19(1):39-44. [doi: 10.1007/s12130-006-1013-7]
54. Kang H, Shin W, Tam L. Differential responses of loyal versus habitual consumers towards mobile site personalization on privacy management. Computers in Human Behavior 2016 Mar;56:281-288. [doi: 10.1016/j.chb.2015.11.013]
55. Ho SY. The Attraction of Internet Personalization to Web Users. Electronic Markets 2006 Feb 1;16(1):41-50. [doi: 10.1080/10196780500491162]
56. van Lier HG, Pieterse ME, Garde A, Postel MG, de Haan HA, Vollenbroek-Hutten MM, et al. A standardized validity assessment protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor. Behav Res Methods 2020 Apr;52(2):607-629 [FREE Full text] [doi: 10.3758/s13428-019-01263-9] [Medline: 31290128]

57. van Manen M. But Is It Phenomenology? Qual Health Res 2017 May;27(6):775-779. [doi: 10.1177/1049732317699570] [Medline: 28682717]

Edited by G Eysenbach; submitted 09.11.19; peer-reviewed by M Johnson, V Nittas; comments to author 05.01.20; revised version received 29.02.20; accepted 20.03.20; published 03.11.20.

Please cite as:
Nelson EC, Sools AM, Vollenbroek-Hutten MMR, Verhagen T, Noordzij ML
Embodiment of Wearable Technology: Qualitative Longitudinal Study
JMIR Mhealth Uhealth 2020;8(11):e16973
URL: https://mhealth.jmir.org/2020/11/e16973
doi:10.2196/16973
PMID:33141093

Review

Wearable Health Devices in Health Care: Narrative Systematic Review

Lin Lu[1*], MD; Jiayao Zhang[1*], MD; Yi Xie[1*], MD; Fei Gao[1], MD; Song Xu[1], MD; Xinghuo Wu[1], MD; Zhewei Ye[1], Prof Dr, MD

Department of Orthopaedic Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[*]these authors contributed equally

Corresponding Author:
Zhewei Ye, Prof Dr, MD
Department of Orthopaedic Surgery
Union Hospital, Tongji Medical College
Huazhong University of Science and Technology
1277 Jiefang Dadao
Wuhan, 430022
China
Phone: 86 17771413685
Email: yezhewei@hust.edu.cn

Abstract

Background: With the rise of mobile medicine, the development of new technologies such as smart sensing, and the popularization of personalized health concepts, the field of smart wearable devices has developed rapidly in recent years. Among them, medical wearable devices have become one of the most promising fields. These intelligent devices not only assist people in pursuing a healthier lifestyle but also provide a constant stream of health care data for disease diagnosis and treatment by actively recording physiological parameters and tracking metabolic status. Therefore, wearable medical devices have the potential to become a mainstay of the future mobile medical market.

Objective: Although previous reviews have discussed consumer trends in wearable electronics and the application of wearable technology in recreational and sporting activities, data on broad clinical usefulness are lacking. We aimed to review the current application of wearable devices in health care while highlighting shortcomings for further research. In addition to daily health and safety monitoring, the focus of our work was mainly on the use of wearable devices in clinical practice.

Methods: We conducted a narrative review of the use of wearable devices in health care settings by searching papers in PubMed, EMBASE, Scopus, and the Cochrane Library published since October 2015. Potentially relevant papers were then compared to determine their relevance and reviewed independently for inclusion.

Results: A total of 82 relevant papers drawn from 960 papers on the subject of wearable devices in health care settings were qualitatively analyzed, and the information was synthesized. Our review shows that the wearable medical devices developed so far have been designed for use on all parts of the human body, including the head, limbs, and torso. These devices can be classified into 4 application areas: (1) health and safety monitoring, (2) chronic disease management, (3) disease diagnosis and treatment, and (4) rehabilitation. However, the wearable medical device industry currently faces several important limitations that prevent further use of wearable technology in medical practice, such as difficulties in achieving user-friendly solutions, security and privacy concerns, the lack of industry standards, and various technical bottlenecks.

Conclusions: We predict that with the development of science and technology and the popularization of personalized health concepts, wearable devices will play a greater role in the field of health care and become better integrated into people's daily lives. However, more research is needed to explore further applications of wearable devices in the medical field. We hope that this review can provide a useful reference for the development of wearable medical devices.

(JMIR Mhealth Uhealth 2020;8(11):e18907) doi:10.2196/18907

KEYWORDS

wearable; medical field; public health; health monitoring; chronic disease management; rehabilitation.

Introduction

Background

In the 1960s, the concept of wearable technology was first proposed by Edward O Thorp [1], a mathematics professor at the Massachusetts Institute of Technology in the United States. Since then, wearable technology has received considerable attention from researchers all around the world. In recent years, with the development of the internet, intelligent hardware, and big data, wearable technology has developed rapidly in various fields such as health care [2], education and culture [3], social networking [4], and the military [5]. Some of these technologies are becoming part of people's daily life in the form of accessories such as smart watches, smart bracelets, armbands, and glasses [6]. In the field of health care, wearable devices in the form of portable medical or health electronic devices that can be directly worn on the body can be used to perceive, record, analyze, regulate, and intervene to maintain health and can even be used to treat diseases with the support of various technologies for identification, sensing, connection, cloud services, and storage [7]. By intelligently integrating mechanical functions with microelectronics and computing power, wearable devices can be used to achieve immediate detection of patient signs and laboratory indicators and provide exercise guidance, drug administration reminders, and so on, with the aim of achieving multiparameter, real-time, online, accurate and intelligent detection and analysis of human physiological and pathological information that can be used to carry out self-diagnosis and self-monitoring [8].

As a standard health care intervention, there are 5 main features of wearable devices [9]: (1) wireless mobility; (2) interactivity and intelligence; (3) sustainability and durability; (4) simple operation and miniaturization; and (5) wearability and portability. From the perspective of modern medicine, the application of wearable devices in the field of medicine follows the 4P medical model characterized by preventive, predictive, personalized, and participatory medicine [10]. On one hand, wearable technologies will play a significant role in advancing precision medicine by enabling measurement of clinically relevant parameters showing the health status of individuals [11]. On the other hand, Loncar-Turukalo et al [12] also indicated that wearable medical devices play an important role as an enabling technology and as a key driver that facilitated the emergence of connected health care. The operation and implementation of these devices depend on the application of various wearable technologies, including sensor technology, medical chip technology, wireless communication technology, power management technology, display technology, and information feedback technology [13]. Real-time medical data from these devices are transmitted to the internet for further analyses or feedback from a health care provider.

Objectives

The development of wearable sensors in the health care market has been relatively slow, despite wearable devices having emerged as a major part of lifestyle and fitness markets. However, the advancement of wearable sensor technology provides enormous opportunities for deployment in health care, especially in connected health care and precision medicine, in which wearable devices can achieve high-quality, real-time measurement of personal health. Although previous reviews have discussed consumer trends in wearable electronics and the application of wearable technology in recreational and sporting activities, data on broad clinical usefulness are lacking. This study reviews the current application of wearable devices in health care. In addition to daily health and safety monitoring, the focus of our work is mainly on the use of wearable devices in clinical practice. We also emphasize their current shortcomings and suggest directions for further research.

Methods

Design

A systematic review design with narrative methods was used to analyze the existing evidence. Specifically, a review methodology [14] was carried out to clarify the types of wearable devices and the current status of their use in health care settings.

Search Strategy

We conducted a comprehensive literature search on January 2, 2020. The following electronic databases were searched with the assistance of an information specialist at the medical library: PubMed, EMBASE, Scopus, and the Cochrane Library. The review was limited to texts published in English between 2015 and 2019 for which abstracts were available. These publication years were chosen due to a dramatic improvement in information technology during that period. The review was also limited to studies of wearable devices in the health care domain. The initial search terms were the following: *wearable devices AND health care*, *wearable technology AND health care*, *sensor AND wearable AND health care*, and *wearable AND track AND health care*. After reviewing the literature identified through these search terms, we added the search terms *monitoring*, *diseases management*, *diagnosis OR treatment*, and *rehabilitation* to capture relevant studies found in the references of the papers retrieved from the initial search.

Inclusion and Exclusion Criteria

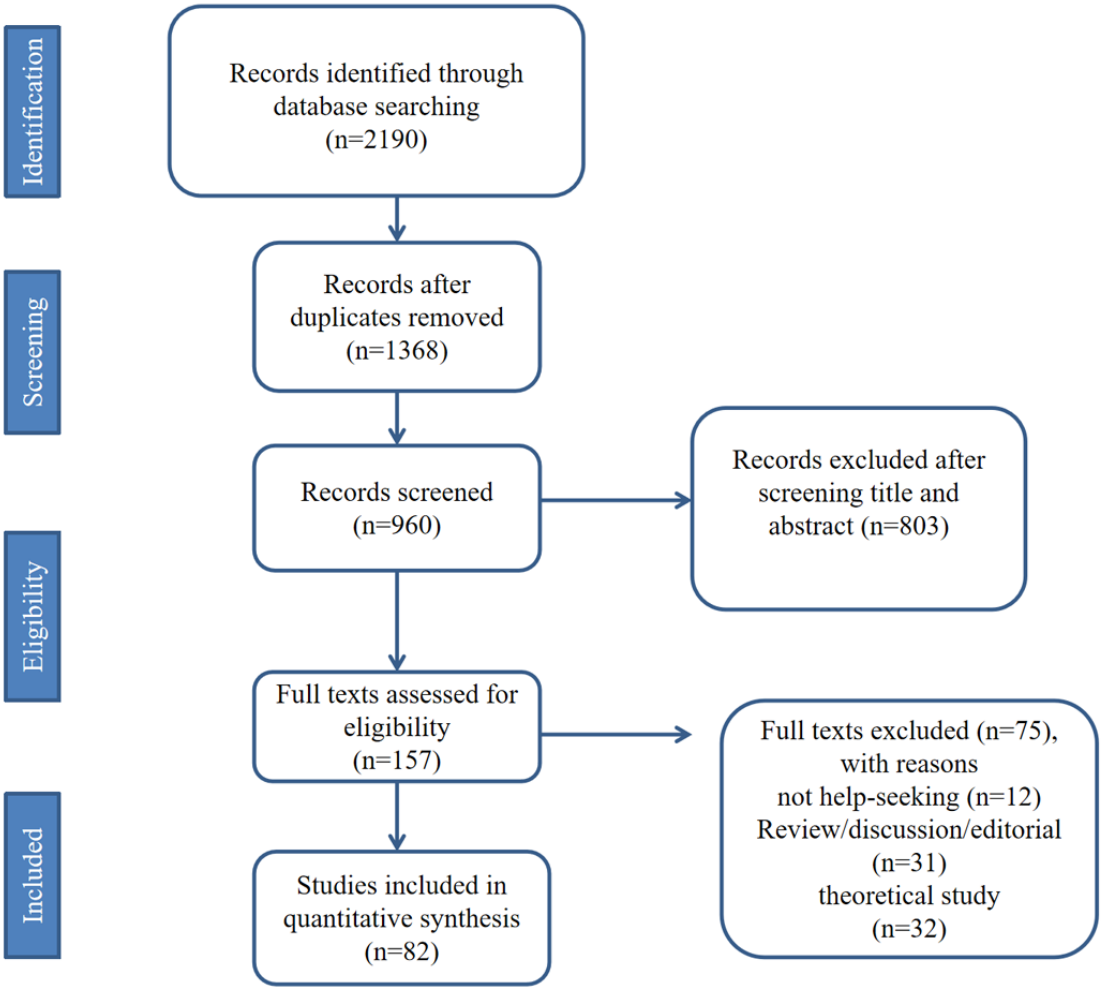
A total of 960 search results were screened for relevance using titles and abstracts, and 82 papers were fully reviewed and are discussed in this study.

Inclusion research criteria were (1) trials including randomized clinical trials and quasi-experimental studies that have proven the effectiveness of wearable devices; (2) studies focusing on clinical applicability; (3) studies published in peer-reviewed journal in English; (4) studies describing completed research; and (5) studies described in full-text papers. There was no restriction on the location of the studies; therefore, international studies written in English were eligible. Exclusion criteria were papers describing the process of the wearable device design, theoretical papers, books or book chapters, letters, statistical reviews, perspectives, dissertations, editorials, and study protocols.

Study Selection

The research screening process consisted of 4 steps. First, 3 authors (LL, ZJY, and XY) independently screened all titles and abstracts of relevance for this systematic review. Second, the abstracts of all relevant papers were screened for eligibility by the 3 authors. Third, the full texts of eligible publications were obtained and screened (LL) according to the inclusion and exclusion criteria. Differences in opinion were resolved by discussing until a consensus was reached. Finally, to avoid incomplete searching, the references of recent related reviews and the primary studies were manually screened for eligible studies.

Figure 1. The flow diagram illustrating the screening process of papers.



Results

After reviewing the title and abstract, the search identified 960 potentially relevant documents. Among them, 82 papers met the inclusion criteria after full-text review (Figure 1).

Classification of Wearable Devices

Wearable devices developed so far have been designed for use on all parts of the human body and are classified into 3 categories: head, limb, and torso wearable devices [15].

Head wearable devices mainly include glasses, helmets, headbands, hearing aids, earrings, earphones, and patches [12]. Among them, Google Glass is a representative type of smart glasses, which has functions such as taking photos, video calls, and GPS positioning, while the application of virtual reality (VR), augmented reality, and mixed reality technology make it suitable for use in areas such as telemedicine, medical education, and intraoperative navigation [16].

Limb wearable devices mainly include wearable devices worn on the arms, legs, and feet [17,18]. Most of the devices worn on the upper limbs are smart watches, bracelets, and other accessories that can be used to monitor physiological parameters such as body temperature, heart rate, ultraviolet exposure levels, and daily activities [19]. Most of the devices worn on the lower limbs appear in the form of shoes and socks that monitor movement-related parameters and are mainly used in the field of rehabilitation [20].

Torso wearable devices mainly include suits, belts, and underwear [21,22]. In recent years, the rapid development of material technology and sensing technology has facilitated the manufacture of electronic products embedded in fabrics or other fabrics that can be used in various biomedical applications. In

2009, a system of clothing providing access to the internet was developed by the Media Technology Laboratory of the Massachusetts Institute of Technology and marked the arrival of a new era of electronic textiles [23].

Application of Wearable Devices in the Medical Field

In the medical field, wearable devices can connect doctors, patients, clouds, and other parties to understand changes in conditions, to alleviate pain, treat diseases, and facilitate the collection of a large sample of case data, which is helpful for the development of national epidemiology strategies and preventive medicine [24]. These devices are used mainly in health and safety monitoring [25], chronic disease management [26], disease diagnosis and treatment [27], and rehabilitation [28].

Health and Safety Monitoring

The health and safety monitoring function of wearable devices is mainly used for older adults, children, pregnant women, and patient groups. The wearer's gait, walking speed, posture, respiratory rate, blood oxygen, heart rate, blood pressure, energy expenditure, position, and other related parameters are monitored in real time to inform nursing requirements [29]. For older adults, a high-quality independent life requires solutions to complex nursing needs related to mobility, intelligence, and independent living, which can be provided by the characteristics of wearable technology [30]. Godfrey et al [31] used wearable devices for gait and fall quantification in older adults, monitoring the viability of older adults' daily activities in an unattended home environment and recognizing the main types of movement (walking, standing, sitting, lying) to help older adults to live independently. Jung et al [32] developed a wearable fall detection system to detect falls by rapidly uploading data for the position of older adult individuals to the medical center and ensuring timely help and treatment. For children, in addition to detecting routine vital signs for health management, wearable technology is a useful tool for tracking children's daily activities [33]. With the rise of wearable devices, children's smart watches, bracelets, and backpacks with tracking and positioning functions have emerged in the market. These devices are based on GPS positioning technology to monitor the child's location and amount of exercise in real time [34]. Furthermore, wearable technology can also be used to monitor mood. Sequeira et al [35] have demonstrated the feasibility of wearable tools in the prediction of depressive symptoms in children and adolescents. Health monitoring of pregnant women can be divided into 2 aspects [36]. On one hand, these applications can be used to monitor the physiological status, emotional state, sleep, and other data before and after conception, while on the other hand, devices can be used to provide immediate feedback on specific problems that occur in the process of child-rearing. Head-med's Compass Pregnancy Monitor [37] is the first medical-grade fetal health home monitoring product that is easy to operate and monitors maternal heart rate, fetal heart rate, and uterine activity through a disposable abdominal patch. For patient groups, wearable devices can be used to monitor symptom changes during treatment, which can be used for disease monitoring and efficacy evaluations and can contribute to individualization of the treatment plan [29]. For example, wearable sensors can be used to monitor the symptoms of patients with Parkinson disease during drug treatment to help doctors adjust drug doses and evaluate the efficacy of new drugs [38]. van Andel et al [39] used photoplethysmography in heart rate monitoring of patients with epilepsy and showed excellent seizure detection performance. Ryvlin et al [40] developed a wearable device that reliably detects generalized tonic-clonic seizures with high sensitivity and specificity to help physicians optimize antiepileptic treatment and reduce the risk of sudden unexpected death in epilepsy.

Chronic Disease Management

Overview

Chronic disease management involves changing passive disease treatment into active health monitoring [41]. Wearable products facilitate data collection and monitoring throughout the user's entire day, as well as providing dynamic, intelligent, and comprehensive analysis of various indicators to enable medical treatment of chronically ill patients. This technology also facilitates remote monitoring of diseases, adjustment of remote treatment plans, lifestyle management, and other functions through cloud services, which is of great significance in disease control [42].

Cardiovascular Disease

Cardiovascular diseases can be easily overlooked, with sudden and potentially lethal consequences that require emergency treatment [43]. They are often accompanied by changes in myocardial electrophysiology at an early stage [44]. Therefore, improving daily monitoring is important for discovering and controlling heart disease. There are 2 kinds of traditional cardiovascular disease monitoring: invasive and noninvasive. For routine monitoring, noninvasive electrocardiograms (ECG) and Doppler echocardiography are the main means of examining cardiac function [45]. The 24-hour ambulatory ECG (Holter monitor), which is a relatively mature wearable medical device currently used in the clinic, allows dynamic monitoring, which conventional ECG does not. However, due to suboptimal wearing comfort and the fact that the conductive gel used with electrodes can lead to chest skin allergies and ulcers, its application in home-based daily monitoring is limited. Ultrasonic heart examination can only be used in a hospital setting [46]. To allow people to manage their own health, researchers at home and abroad have conducted many studies on wearable health monitoring systems, especially for ECG long-term data collection. Tsukada et al [47] developed a sports vest made from nanofibers coated with an electroconductive polymer to place the ECG electrodes in close contact with the human body. The signal allows display of the ECG signal in real time, and the monitoring data are collected through an app, which increases comfort without risking allergies. The data are uploaded to the cloud and analyzed by a professional physician to realize remote monitoring of heart diseases. In December 2018, Apple developed the Apple Watch Series 4, which combined the functions of ECG and a watch for the first time. In addition, the dial was designed to display a bipolar ECG to monitor occult atrial fibrillation. This device shows similar accuracy in monitoring arrhythmia, atrioventricular block, and

QRS duration extension to standard 12-lead ECG recordings [48]. Kaspar et al [49] found that the use of a wearable cardioverter-defibrillator protected patients from sudden cardiac arrest when they were treated in nonhospital settings until reimplantation of an implantable cardioverter-defibrillator.

Pulmonary Diseases

Acute exacerbations of chronic obstructive pulmonary disease (COPD) and bronchial asthma can lead to impaired lung function, decreased quality of life, and increased mortality [50]. Active monitoring of the early signs of a patient's condition and early treatment can prevent these outcomes. The telehealth program aims to facilitate early identification and timely self-management of acute exacerbations of COPD and bronchial asthma [51]. For such patients, early detection of progression can help to control the disease [52]. The emergence of inexpensive wearable devices has enabled people to monitor heart rate, pulse, oxygen saturation, and physical activity continuously, as well as audio to detect cough, breath sounds, and other characteristics [53]. These signals can be used in predictive analyses to detect early deterioration of lung function. A prospective cohort study [54] conducted at the University of Toronto evaluated a wearable system that reliably captured nearly continuous patient respiratory rate, oxygen saturation, heart rate, and other data for screening of early COPD deterioration. The results of this study demonstrated the feasibility of using a smart watch for centralized monitoring of patients with COPD. A wearable device developed by Colantonio et al [55] uses a wireless sensor network system to monitor the patient's respiratory rate, respiratory sound, blood oxygen saturation, and ECG to evaluate the therapeutic effect of treatments for COPD. Li et al [56] used acoustic respirators to monitor nighttime wheezing in asthmatic children and found that 57% of patients with well-controlled asthma had a significant number of nighttime wheezing episodes and poor association with routine measurements of lung function. This has helped to develop individualized treatment for children with asthma. However, Rubio et al [51] reported disappointing results for the use of wearable devices to monitor acute exacerbations of COPD, in part because the parameters monitored (symptoms, pulse oximetry, and lung volume) were not reliable indicators for predicting exacerbation of the disease. It should be pointed out that with advances in sensing technology, wearable systems can also link individual environmental exposure with physiology and subsequent adverse health reactions, providing clues for the pathogenesis of certain lung diseases [57].

Diabetes

Diabetes represents a group of metabolic diseases characterized by hyperglycemia caused by defects in insulin secretion or impaired biological effects [58]. Long-term poor glycemic control can lead to damage, dysfunction, and failure of various organ tissues, especially the eyes, kidneys, nerves, heart, and blood vessels [59]. For patients with diabetes, improving the ability of self-monitoring and self-management of blood glucose levels has contributed to the reduction of diabetes-related morbidity and mortality. There are currently 3 types of medical management products on the market for patients with diabetes: blood glucose level monitoring equipment, injectable insulin, and implantable insulin pumps [60]. Among them, blood glucose level monitoring products have an important position in blood glucose level control, which is the basic reference for the adjustment of other treatment methods and can also prevent the occurrence of risk events [61]. Traditional blood glucose level monitoring is performed by directly drawing a venous blood sample or taking a finger-prick blood sample, which is analyzed by a biochemical analyzer. These methods are invasive and inconvenient, especially for patients with diabetes who need to monitor blood glucose level several times a day. Due to the volatility and transience of blood glucose level testing, traditional single-point testing methods do not truly reflect the changes in glucose levels in the body [62]. With the development of mobile and sensor technologies, wearable dynamic blood glucose level monitoring products have emerged. The GlucoWatch [61], a noninvasive, painless blood glucose level monitoring product approved by the US Food and Drug Administration in March 2001, has proven its applicability and feasibility in the field of diabetes monitoring products. At present, wearable medical devices widely use indirect measurement methods (minimally invasive or noninvasive) to measure parameters such as blood glucose concentration. The main methods are spectrometry, blood substitution (urine, tears, and tissue fluid), counter-ion electroosmosis, and microwave technology [63,64]. Compared with other methods, the optical method is rapid, is noninvasive, is nonpolluting, is simple to operate, and has become the main method for noninvasive blood glucose level detection. The principle of the measurement is based on concentration-dependent changes in the absorption and reflection characteristics of glucose [65]. However, the accuracy of the measurement is limited by the overlap of other blood components absorption spectra with the absorption spectrum of glucose [64]. In addition, Medtronic's Minimed 670G uses a portable design to integrate blood glucose level monitoring and an insulin pump. The patient can affix the device to the waist and set the daily detection and injection times. Pillalamarri et al [66] developed a handheld insulin pump using biomedical microelectromechanical systems technology to intelligently control the rate and volume of insulin injections, maintaining blood glucose level within a relatively stable range.

Hypertension

Hypertension is a chronic disease characterized by a sustained increase in arterial blood pressure. It is the most important risk factor for cardiovascular and cerebrovascular diseases, which seriously endanger human health [67]. Therefore, the accuracy and reliability of blood pressure measurement during the diagnosis and treatment of hypertension is especially important [68]. Blood pressure can be measured directly and indirectly [69]. Direct measurement involves percutaneous puncture and catheterization of the aorta. This is an invasive method and is only applicable in critical and difficult cases. Indirect measurement, also known as cuff compression, involves the use of a sphygmomanometer. This is the most commonly used measurement method, but continuous data cannot be obtained [70]. Noninvasive continuous blood pressure measurement is the trend for the development of wearable blood pressure monitors. At present, wearable devices determine blood pressure by measuring different physiological signals and can be divided

into 4 types according to the principle [71]: (1) flattening tension of the radial artery; (2) volume changes in pulsing blood; (3) speed of pulse wave; and (4) the vibration measurement method. Moreover, wearable blood pressure monitor can be roughly divided into 2 types according to structure: cuff type and sleeveless type. Due to its strong anti-interference and reliability, the cuff type has become the mainstream form of wearable blood pressure measuring devices and are widely used in the market [72]. This type of wearable device uploads the monitored data to generate a dynamic blood pressure map, which is convenient for patients and doctors. However, during daily continuous monitoring, repeated inflation and deflation of the cuff can cause physical discomfort to the patient, especially at night when the process can cause sleep disturbances [70]. Therefore, a wearable device that can measure dynamic blood pressure accurately and comfortably without a cuff would be attractive prospect. Zheng et al [73] used a wearable sleeveless device developed based on optical technology to monitor blood pressure changes by measuring the pulse arrival time (the pulse transit time from the heart to the peripheral blood vessels). In a mixed methods study, Islam et al [74] demonstrated that wearable blood pressure device measurements compared well against a gold-standard ambulatory device, indicating that this user-friendly method has the potential to enhance blood pressure management in long-term monitoring. However, this type of sleeveless equipment is still in the experimental development stage and has not yet entered the market.

Diagnosis and Treatment of Diseases

Overview

A comprehensive understanding of the changes in physiological and pathological indicators during the early stage of disease is critical for timely diagnoses and interventions. Wearable devices are of great significance in the diagnosis and treatment of various diseases by monitoring the changes in vital signs on a real-time basis [42].

Neurological Disorders

For example, early warning and intervention in the predementia phase of Alzheimer disease are of great significance in delaying the onset and reducing the incidence [75]. Mild cognitive dysfunction is a major feature of predementia in Alzheimer disease, and diagnostic methods at this stage are not yet fully developed [76]. Recent studies have shown that gait is a noninvasive biological indicator of cognitive function [75,77,78]. By wearing a wearable device, the user's gait parameters can be collected for early detection of Alzheimer disease. In addition, wearable devices also show good application prospects in the early diagnosis of other neurological diseases. For instance, Mannini et al [79] developed a wearable device that analyzes gait classification to improve the accuracy of diagnosis of early neurological diseases.

Respiratory Diseases

For patients with obstructive sleep apnea hypopnea syndrome, the application of wearable nocturnal breathing monitoring equipment can improve the accuracy of early diagnosis and can be used at home [80]. Surrel et al [81] developed a wearable, accurate, and energy efficient system for monitoring obstructive sleep apnea on a long-term basis.

Cardiac Conduction System Anomalies

For patients who may be at risk of cardiac arrest, a wearable defibrillator can help monitor arrhythmias. Moreover, emergency defibrillation can be performed to restore normal rhythm when cardiac arrest or ventricular fibrillation is detected [82]. The annual meeting of the American Association for Cancer Research in 2017 also reported that a medical wearable device that delivers alternating currents extends the overall survival of patients with malignant gliomas [83].

Urinary Diseases

For end-stage renal disease, Gura et al [84] showed that treatment with a wearable artificial kidney was well tolerated and resulted in effective uremic solute clearance and maintenance of electrolyte and fluid homeostasis.

However, wearable devices still have a long way to go in terms of therapeutic applications compared to their health monitoring functions. It should be pointed out that in recent years, with the rise of VR, augmented reality, and mixed reality technology and the breakthrough of remote technology, wearable equipment has also undergone significant developments for application in settings such as medical education, the formulation of preoperative surgery plans, intraoperative navigation, preoperative doctor-patient communication, and remote consultation [16]. In our previous studies [16,85-87], we successfully implemented mixed reality technology to facilitate preoperative communications between doctors and patients, intraoperative guidance, remote surgical consultation, and surgical navigation through HoloLens glasses.

Rehabilitation

Overview

In the field of rehabilitation, wearable devices are used mainly in sports rehabilitation [88], cognitive rehabilitation [89], and as rehabilitation aids for people with disabilities [90].

Sports Rehabilitation

Conditions such as stroke, brain trauma, spinal cord injury, and musculoskeletal injury often lead to the loss or decrease of a patient's motor ability. The main goal of sports rehabilitation is to restore balance and coordination, to ensure normal joint mobility, and to have sufficient muscle strength and muscular endurance for daily life [91]. Traditional sports rehabilitation is mainly performed in specific medical institutions by professional rehabilitation practitioners who carry out training to expand the range of motion of joints, enhance muscle strength and endurance, and improve balance and coordination function. This mode of rehabilitation training has the advantages of safe reliable methods and real-time guidance from professionals [92]. At the same time, there are some shortcomings in the traditional rehabilitation model that cannot be ignored, such as limitations in the time and place of rehabilitation, as well as the boredom and tedium of the process, resulting in lack of adherence among patients, even those who are not older adults, all of which seriously affect the efficacy of sports rehabilitation [21]. The application of wearable devices, especially in

combination with VR, augmented reality, and mixed reality technologies, not only offers the ability to comprehensively monitor and evaluate the rehabilitation activities of patients but also to make the activities more interesting and improve patient adherence [92].

Limb hemiplegia after brain injury is a difficult problem in sports rehabilitation [93]. According to The Chinese Stroke Prevention Report in 2018, stroke has become the leading cause of death in China and the leading cause of disability in Chinese adults [94]. In an aging society, the incidence of stroke is expected to increase in the coming years. Among the various defects caused by stroke, unilateral sensorimotor deficits are very prominent, and 80% of stroke patients have different degrees of gait abnormalities [95]. At present, lower limb rehabilitation for stroke patients is focused mainly on gait training [28]. The wearable device can be used to monitor the patient's gait parameters and provide feedback to help the doctor assess the patient's recovery in real time so that the treatment plan can be adjusted accordingly. Hsu et al [96] used a multiplaced wearable sensor to analyze and classify gait characteristics in patients with neurological disorders to guide the selection of rehabilitation exercise regimens. Furthermore, in cases of hemiplegia of a single limb caused by neurological disease, the recovery of arm function lags behind other functions, although the posture and gait may be significantly improved [97]. Maceira-Elvira et al [98] used a wearable stroke rehabilitation trainer to support patients performing a personalized upper limb neuromotor training program at home. The built-in wireless sensor records the patient's exercise volume, analyzes the data and feeds the data back to the patient and therapist, thus bringing additional improvement to the recovery of the patient's upper limb functions. Using an optimized wearable glove, a team from the University of Pisa used hand posture reconstruction technology to reconstruct hand movements, allowing real-time recording and feedback on a patient's hand function recovery [99].

Moreover, in a broad sense, the rehabilitation treatment of spinal cord injury and musculoskeletal injury can be categorized as orthopedic rehabilitation. The use of orthopedic aids is particularly important in the treatment of modern orthopedic rehabilitation. For patients with spinal cord injury, early rehabilitation treatment leads to better the recovery of spinal nerve function mainly through improvements in nerve plasticity [100]. Traditional rehabilitation treatment, which is guided mainly by professional rehabilitators in hospitals, is a costly process requiring a long period of hospitalization. Furthermore, the qualification requirements of the rehabilitation professionals are often high, and one-on-one professional guidance is not always possible. Many researchers [101,102] have designed wearable artificial exoskeletons based on bionics, the principle of which is to transfer the human body load to the artificial exoskeleton and assist in maintaining the standing posture of patients with spinal cord injury. This approach is also designed to allow the patient to walk slowly with a fixed gait maintained by the dynamics and feedback systems of the artificial exoskeleton, thus avoiding joint dysfunction such as joint stiffness and muscle contracture [103]. On the other hand, 3 factors are key to the success of treatment of musculoskeletal system injury: reduction, fixation, and functional exercise [104]. The rehabilitation exercises in the later stage after the injury are the most important. Many orthopedic surgeons in clinical practice often consider only the aesthetics of the operation, neglecting the importance of postoperative limb function rehabilitation. Furthermore, many patients pay insufficient attention to postoperative rehabilitation exercise, with poor adherence, and fear of pain, all of which affect the recovery of the function of the affected limb and greatly reduce the effectiveness of the surgery. The use of wearable devices can encourage patients to perform rehabilitation exercises and allow adjustment of the intensity of rehabilitation training to improve training results based on feedback information [103]. The electronic-assisted instrument system developed by Zhu et al [105] promoted the recovery of knee flexion and extension function in patients after total knee arthroplasty. Lee et al [106] designed an exoskeleton suit that can assist multiple joint movements and measure the direction and angle of movement. Through intuitive recording of the data, which is displayed in a graphical form, the exoskeleton suit provides patients and physicians with information about the effectiveness and extent of the exercise, which is conducive to early rehabilitation and restoration of function.

Cognitive Rehabilitation

Cognitive dysfunction, which is one of the most common sequelae after brain injury, not only influences the quality of life of patients but also puts tremendous pressure on patients' families and society [107]. Therefore, improvements in rehabilitation methods are of great significance for cognitive dysfunction. VR glasses have shown great potential in providing treatment options and assessment tools for patients with cognitive impairment [108-110]. VR technology has 3 characteristics, namely immersion, interactivity, and imagination [111]. By providing visual, auditory, and tactile sensory simulation, patients are provided with an immersive experience that aids cognitive rehabilitation in a controlled stimulation environment and facilitates monitoring of related parameters in real time. VR technology offers the potential to develop a personalized treatment plan for patients with different levels of cognitive impairment by providing virtual reproducible images, which are effective in the recovery of memory impairment. To date, many studies have demonstrated the value of VR-based wearable devices in the rehabilitation of cognitive impairment. According to the characteristics of immersion, Faria et al [112] designed VR glasses to perform memory recovery training in patients with cognitive dysfunction, reducing their fear of reality and improving their learning and behavioral abilities. Wåhlin et al [113] used VR training to improve left-sided awareness in chronic stroke patients also increased sporadic interhemispheric functional connectivity within the dorsal attention network.

Rehabilitation Aids for People With Disabilities

Accessories for people with disabilities is another major direction in the field of wearable devices for rehabilitation. Such accessories include intelligent glasses for the blind, smart hearing aids, and intelligent prosthetics.

Using artificial intelligence technology, smart glasses for the blind [114] not only help individuals with visual impairments

interpret information about their surroundings (identified by a pinhole camera, eg, to analyze road condition information in real time) but also aid in making effective decisions. Moreover, smart glasses can help individuals with visual impairment to use a variety of smart home products in a family setting, thus allowing them to improve their quality of life [115].

The hearing aid is essentially an electroacoustic amplifier [116]. The acoustic signal is converted into an electrical signal by a microphone, and after amplification, the electrical signal is restored by the receiver to an acoustic signal and transmitted to the human ear. The emergence of intelligent technology has promoted the development of the hearing aid industry, allowing users to autonomously choose the clearest sound they want to hear in a complex and changing environment, freeing them from the disadvantages of hearing disorders.

For patients with limb defects, the prosthesis can not only fill the shape defect but also restore full or partial residual limb function [117]. Traditional prosthetics are cumbersome, and the invention of wearable smart prostheses is very encouraging for patients with physical disabilities. Intelligent prostheses incorporating robotics have become a hot topic in recent years [118].

Discussion

Existing Issues

Overview

The development of wearable medical equipment has increased the popularity and quality of health care. Moreover, these developments have, to a certain extent, alleviated the shortage of medical resources in low-income countries and promoted the development of medical care worldwide. However, the wearable medical device industry is still in development and currently faces several important limitations that prevent further use of wearable technology in medical practice.

Barriers to User-Friendly Solutions

For the health care system, the main challenge is to enable the use of these technologies by changing the model of care and by sharing information [119]. Data collection, transfer, preservation, and sharing require not only the development of technical solutions but also the development of legal infrastructure, which will enable different organizations to share data and responsibilities for patient [120]. On one hand, patient autonomy in using wearable health devices needs to be kept to help patients to become active participants in their own care [119]. On the other hand, the duty of clinicians and the responsibilities for misdiagnosis and missed diagnosis due to the use of unreliable or delayed data and false alarms during the use of wearable devices should also be legally regulated [121].

Security and Privacy Concerns

Through sensor technology, wearable health devices can collect all kinds of user information, such as health information, geographical location, and living habits [34]. The various formats, large scale, and numerous mobile links of these data may increase the risk of leakage and tampering [8,35,122,123]. Strategies to ensure the security of the data and improve the public trust are required.

Lack of Industry Standards and Regulations

In the absence of industry standards and regulations, all companies hope to rely on their core products to form their own standards and regulations, making it difficult to integrate resources. So, the establishment and enforcement of new regulatory standards are required [71,118].

Technological Barriers

There are many technical bottlenecks in the applications of wearable medical devices [64,124,125]: (1) Data accuracy [9,24,118,122]: On one hand, the sensor specificity of current wearable device is low, which may lead to overdetection of benign nonclinical related signals, resulting in misdiagnosis, unnecessary examinations, and patient anxiety. On the other hand, low sensor sensitivity may lead to omission of pathological clinically-relevant parameters, resulting in missed diagnosis and delay in treatment. (2) Single function: The compatibility of wearable devices is relatively poor, the functions are mainly concentrated on the level of health monitoring, with slow progress in clinical treatment, and few wearable medical devices have effectively integrated multiple functions [27,124]. (3) Poor battery life: Designing low-power consuming and high-energy storage wearable devices has always been an exciting but challenging issue [10,120]. (4) Equipment safety: Safety and security are primary considerations for medical equipment, closely related to reliability at all system levels [12,13]. Because false alarms would reduce user alertness and prevent user adherence from the feedback provided, efficient mechanisms are needed to detect and diagnose deviations occurring in captured data [126]. (5) Other issues such as cost, low data collection and processing efficiency, unstable human-computer interaction interfaces, and incomplete construction of big data health clouds need to be further improved [9,13,122].

Limitations

There were some limitations in our review. Although some databases are included, our search terms may cause the omission of relevant papers. Due to the exploratory nature of this review, it included a wide range of research designs, and the review will ultimately be limited by the design of the included studies. Although the second and third reviewers used strategies to limit bias through consultation, they also acknowledge the possibility of subjectivity in analyzing the survey results. Additionally, this paper is not intended to be a systematic review, and it is possible to conduct a broader review and find papers suggesting other applications of wearable devices.

Conclusion and Future Directions

Despite many limitations in their application, wearable devices have achieved remarkable success and have brought huge benefits to the aging global population. The authors believe that the following aspects are crucial for future development of the wearable device industry:

1. Strengthening oversight of the wearable device industry, formulating specific security rules to protect the privacy

and security of personal data, and clarifying the relevant medical responsibilities and rights between doctors and patients.

2. Establishment and enforcement of industry standards by taking health care as the main body of information to establish a unified data classification, evaluation system, and industry standards for wearable device to realize the mutual recognition of medical data.
3. Technological advancement in wearable health devices to develop low-consumption and high-integration sensor technology; low-power high-performance battery technology; high processing efficiency medical chip technology; and human-computer interaction technology to improve information accuracy, information processing speed, extend battery life, and user experience.
4. Combining big data, cloud computing, and IoT—the internet of things—to build a healthy database to develop a complete medical ecosystem. Using this resource, we can fully develop, analyze, and employ medical and health big data to expand the use of wearable devices in other fields, such as telemedicine, preventive medicine, and epidemiology [12].

By combining smart wearable medical devices with pension services, a smart retirement community can be built to provide high-quality, high-efficiency medical care services. How medical services are sought has also begun to transform from passive medical treatment of disease to community medical models led by prevention, health care, and prediagnosis.

In summary, with the development of science and technology and the popularization of personalized health concepts, wearable devices will inevitably play a greater role in the field of health care and become better integrated into our daily lives. However, more research is needed to explore further applications of wearable devices in the medical field.

Acknowledgments

This work was supported by Foundation for the Science and Technology Commission of Hubei (grant no. 2018AAA067) and the National Natural Science Foundation of China (no. 81974355).

Conflicts of Interest

None declared.

References

1. Amft O, Lukowicz P. From backpacks to smartphones: past, present, and future of wearable computers. IEEE Pervasive Comput 2009 Jul;8(3):8-13 [FREE Full text] [doi: 10.1109/MPRV.2009.44]
2. Khan Y, Ostfeld AE, Lochner CM, Pierre A, Arias AC. Monitoring of vital signs with flexible and wearable medical devices. Adv Mater 2016 Jun;28(22):4373-4395 [FREE Full text] [doi: 10.1002/adma.201504366] [Medline: 26867696]
3. Guze PA. Using technology to meet the challenges of medical education. Trans Am Clin Climatol Assoc 2015;126:260-270 [FREE Full text] [Medline: 26330687]
4. Jackson MM, Valentin G, Freil L, Burkeen L, Zeagler C, Gilliland S, et al. FIDO—Facilitating interactions for dogs with occupations: wearable communication interfaces for working dogs. Pers Ubiquit Comput 2015 Oct 11;19(1):155-173 [FREE Full text] [doi: 10.1007/s00779-014-0817-9]
5. Siddall AG, Powell SD, Needham-Beck SC, Edwards VC, Thompson JE, Kefyalew SS, et al. Validity of energy expenditure estimation methods during 10 days of military training. Scand J Med Sci Sports 2019 Sep;29(9):1313-1321 [FREE Full text] [doi: 10.1111/sms.13488] [Medline: 31136027]
6. Park Y, Lee S, Park J. Recent progress in wireless sensors for wearable electronics. Sensors (Basel) 2019 Oct 09;19(20) [FREE Full text] [doi: 10.3390/s19204353] [Medline: 31600870]
7. Haghi M, Thurow K, Stoll R. Wearable devices in medical internet of things: scientific research and commercially available devices. Healthc Inform Res 2017 Jan;23(1):4-15 [FREE Full text] [doi: 10.4258/hir.2017.23.1.4] [Medline: 28261526]
8. Guk K, Han G, Lim J, Jeong K, Kang T, Lim E, et al. Evolution of wearable devices with real-time disease monitoring for personalized health care. Nanomaterials (Basel) 2019 May 29;9(6) [FREE Full text] [doi: 10.3390/nano9060813] [Medline: 31146479]
9. Qian R, Long Y. Wearable chemosensors: a review of recent progress. Chemistry Open 2018 Feb;7(2):118-130 [FREE Full text] [doi: 10.1002/open.201700159] [Medline: 29435397]
10. Lin B. Wearable smart devices for p4 medicine in heart disease: ready for medical cyber-physical systems? OMICS 2019 May;23(5):291-292 [FREE Full text] [doi: 10.1089/omi.2019.0059] [Medline: 31050597]
11. Jeong IC, Bychkov D, Searson PC. Wearable devices for precision medicine and health state monitoring. IEEE Trans Biomed Eng 2019 May;66(5):1242-1258 [FREE Full text] [doi: 10.1109/tbme.2018.2871638]
12. Loncar-Turukalo T, Zdravevski E, Machado da Silva J, Chouvarda I, Trajkovik V. Literature on wearable technology for connected health: scoping review of research trends, advances, and barriers. J Med Internet Res 2019 Sep 05;21(9):e14017 [FREE Full text] [doi: 10.2196/14017] [Medline: 31489843]
13. Sim I. Mobile devices and health. N Engl J Med 2019 Sep 05;381(10):956-968 [FREE Full text] [doi: 10.1056/NEJMra1806949] [Medline: 31483966]

14. Grant M, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J 2009 Jun;26(2):91-108 [FREE Full text] [doi: 10.1111/j.1471-1842.2009.00848.x] [Medline: 19490148]
15. Koydemir HC, Ozcan A. Wearable and implantable sensors for biomedical applications. Annu Rev Anal Chem (Palo Alto Calif) 2018 Jun 12;11(1):127-146. [doi: 10.1146/annurev-anchem-061417-125956] [Medline: 29490190]
16. Hu H, Feng X, Shao Z, Xie M, Xu S, Wu X, et al. Application and prospect of mixed reality technology in medical field. Curr Med Sci 2019 Feb;39(1):1-6. [doi: 10.1007/s11596-019-1992-8] [Medline: 30868484]
17. Wang Q, Markopoulos P, Yu B, Chen W, Timmermans A. Interactive wearable systems for upper body rehabilitation: a systematic review. J Neuroeng Rehabil 2017 Mar 11;14(1):20 [FREE Full text] [doi: 10.1186/s12984-017-0229-y] [Medline: 28284228]
18. Walmsley CP, Williams SA, Grisbrook T, Elliott C, Imms C, Campbell A. Measurement of upper limb range of motion using wearable sensors: a systematic review. Sports Med Open 2018 Nov 29;4(1):53 [FREE Full text] [doi: 10.1186/s40798-018-0167-7] [Medline: 30499058]
19. 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]
20. Powell L, Parker J, Martyn St-James M, Mawson S. The effectiveness of lower-limb wearable technology for improving activity and participation in adult stroke survivors: a systematic review. J Med Internet Res 2016 Oct 07;18(10):e259 [FREE Full text] [doi: 10.2196/jmir.5891] [Medline: 27717920]
21. Xu J, Bao T, Lee UH, Kinnaird C, Carender W, Huang Y, et al. Configurable, wearable sensing and vibrotactile feedback system for real-time postural balance and gait training: proof-of-concept. J Neuroeng Rehabil 2017 Oct 11;14(1):102 [FREE Full text] [doi: 10.1186/s12984-017-0313-3] [Medline: 29020959]
22. Argent R, Slevin P, Bevilacqua A, Neligan M, Daly A, Caulfield B. Clinician perceptions of a prototype wearable exercise biofeedback system for orthopaedic rehabilitation: a qualitative exploration. BMJ Open 2018 Oct 25;8(10):e026326. [doi: 10.1136/bmjopen-2018-026326] [Medline: 30366919]
23. Zhang H, Liu F. Updated Technology and Application of Intelligent Electric Wears. China Textile Leader 2011;12:108-110 [FREE Full text] [doi: 10.3969/j.issn.1003-3025.2011.12.029]
24. Burmester GR. Rheumatology 4.0: big data, wearables and diagnosis by computer. Ann Rheum Dis 2018 Jul;77(7):963-965. [doi: 10.1136/annrheumdis-2017-212888] [Medline: 29802224]
25. Jo A, Coronel BD, Coakes CE, Mainous AG. Is there a benefit to patients using wearable devices such as Fitbit or health apps on mobiles?: a systematic review. Am J Med 2019 Dec;132(12):1394-1400.e1 [FREE Full text] [doi: 10.1016/j.amjmed.2019.06.018] [Medline: 31302077]
26. Maddison R, Cartledge S, Rogerson M, Goedhart NS, Ragbir Singh T, Neil C, et al. Usefulness of wearable cameras as a tool to enhance chronic disease self-management: scoping review. JMIR Mhealth Uhealth 2019 Jan 03;7(1):e10371 [FREE Full text] [doi: 10.2196/10371] [Medline: 30609985]
27. Ray PP, Dash D, De D. A systematic review of wearable systems for cancer detection: current state and challenges. J Med Syst 2017 Oct 02;41(11):180 [FREE Full text] [doi: 10.1007/s10916-017-0828-y] [Medline: 28971278]
28. Jayaraman C, Mummidisetty CK, Mannix-Slobig A, McGee Koch L, Jayaraman A. Variables influencing wearable sensor outcome estimates in individuals with stroke and incomplete spinal cord injury: a pilot investigation validating two research grade sensors. J Neuroeng Rehabil 2018 Mar 13;15(1):19 [FREE Full text] [doi: 10.1186/s12984-018-0358-y] [Medline: 29534737]
29. Takei K, Honda W, Harada S, Arie T, Akita S. Toward flexible and wearable human-interactive health-monitoring devices. Adv Healthc Mater 2015 Mar 11;4(4):487-500 [FREE Full text] [doi: 10.1002/adhm.201400546] [Medline: 25425072]
30. 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 [FREE Full text] [doi: 10.1007/s10916-019-1365-7] [Medline: 31203472]
31. Godfrey A. Wearables for independent living in older adults: Gait and falls. Maturitas 2017 Jun;100:16-26 [FREE Full text] [doi: 10.1016/j.maturitas.2017.03.317] [Medline: 28539173]
32. Jung S, Hong S, Kim J, Lee S, Hyeon T, Lee M, et al. Wearable fall detector using integrated sensors and energy devices. Sci Rep 2015 Nov 24;5:17081 [FREE Full text] [doi: 10.1038/srep17081] [Medline: 26597423]
33. Mackintosh KA, Chappel SE, Salmon J, Timperio A, Ball K, Brown H, et al. Parental perspectives of a wearable activity tracker for children younger than 13 years: acceptability and usability study. JMIR Mhealth Uhealth 2019 Nov 04;7(11):e13858 [FREE Full text] [doi: 10.2196/13858] [Medline: 31682585]
34. Wen D, Zhang X, Lei J. Consumers' perceived attitudes to wearable devices in health monitoring in China: A survey study. Comput Methods Programs Biomed 2017 Mar;140:131-137 [FREE Full text] [doi: 10.1016/j.cmpb.2016.12.009] [Medline: 28254069]
35. Sequeira L, Perrotta S, LaGrassa J, Merikangas K, Kreindler D, Kundur D, et al. Mobile and wearable technology for monitoring depressive symptoms in children and adolescents: A scoping review. J Affect Disord 2020 Mar 15;265:314-324 [FREE Full text] [doi: 10.1016/j.jad.2019.11.156] [Medline: 32090755]

36. Prioreschi A, Nappey T, Westgate K, Olivier P, Brage S, Micklesfield LK. Development and feasibility of a wearable infant wrist band for the objective measurement of physical activity using accelerometery. Pilot Feasibility Stud 2018;4:60 [FREE Full text] [doi: 10.1186/s40814-018-0256-x] [Medline: 29507750]
37. Bachiri M, Idri A, Fernández-Alemán JL, Toval A. Mobile personal health records for pregnancy monitoring functionalities: Analysis and potential. Comput Methods Programs Biomed 2016 Oct;134:121-135 [FREE Full text] [doi: 10.1016/j.cmpb.2016.06.008] [Medline: 27480737]
38. Godinho C, Domingos J, Cunha G, Santos AT, Fernandes RM, Abreu D, et al. A systematic review of the characteristics and validity of monitoring technologies to assess Parkinson's disease. J Neuroeng Rehabil 2016 Mar 12;13:24 [FREE Full text] [doi: 10.1186/s12984-016-0136-7] [Medline: 26969628]
39. van Andel J, Ungureanu C, Aarts R, Leijten F, Arends J. Using photoplethysmography in heart rate monitoring of patients with epilepsy. Epilepsy Behav 2015 Apr;45:142-145 [FREE Full text] [doi: 10.1016/j.yebeh.2015.02.018] [Medline: 25812938]
40. Ryvlin P, Ciumas C, Wisniewski I, Beniczky S. Wearable devices for sudden unexpected death in epilepsy prevention. Epilepsia 2018 Jun;59 Suppl 1:61-66 [FREE Full text] [doi: 10.1111/epi.14054] [Medline: 29873831]
41. van der Heide I, Poureslami I, Mitic W, Shum J, Rootman I, FitzGerald JM. Health literacy in chronic disease management: a matter of interaction. J Clin Epidemiol 2018 Oct;102:134-138 [FREE Full text] [doi: 10.1016/j.jclinepi.2018.05.010] [Medline: 29793001]
42. Sun J, Guo Y, Wang X, Zeng Q. mHealth for aging China: opportunities and challenges. Aging Dis 2016 Jan;7(1):53-67 [FREE Full text] [doi: 10.14336/AD.2015.1011] [Medline: 26816664]
43. Singh S, Kullo IJ, Pardi DS, Loftus EV. Epidemiology, risk factors and management of cardiovascular diseases in IBD. Nat Rev Gastroenterol Hepatol 2015 Jan;12(1):26-35 [FREE Full text] [doi: 10.1038/nrgastro.2014.202] [Medline: 25446727]
44. Dolabchian ZL. [Electrophysiological classification of myocardial infarct]. Kardiologiia 1975 Jan;15(1):113-115 [FREE Full text] [Medline: 1117614]
45. Fumagalli S, Chen J, Dobreanu D, Madrid AH, Tilz R, Dagres N. The role of the Arrhythmia Team, an integrated, multidisciplinary approach to treatment of patients with cardiac arrhythmias: results of the European Heart Rhythm Association survey. Europace 2016 May 12;18(4):623-627. [doi: 10.1093/europace/euw090]
46. Steinberg C, Philippon F, Sanchez M, Fortier-Poisson P, O'Hara G, Molin F, et al. A novel wearable device for continuous ambulatory ECG recording: proof of concept and assessment of signal quality. Biosensors (Basel) 2019 Jan 21;9(1) [FREE Full text] [doi: 10.3390/bios9010017] [Medline: 30669678]
47. Tsukada YT, Tokita M, Murata H, Hirasawa Y, Yodogawa K, Iwasaki Y, et al. Validation of wearable textile electrodes for ECG monitoring. Heart Vessels 2019 Jul;34(7):1203-1211 [FREE Full text] [doi: 10.1007/s00380-019-01347-8] [Medline: 30680493]
48. Ip JE. Wearable devices for cardiac rhythm diagnosis and management. JAMA 2019 Jan 29;321(4):337-338 [FREE Full text] [doi: 10.1001/jama.2018.20437] [Medline: 30633301]
49. Kaspar G, Sanam K, Gholkar G, Bianco NR, Szymkiewicz S, Shah D. Long-term use of the wearable cardioverter defibrillator in patients with explanted ICD. Int J Cardiol 2018 Dec 01;272:179-184 [FREE Full text] [doi: 10.1016/j.ijcard.2018.08.017] [Medline: 30121177]
50. Homma T, Kawahara T, Mikuni H, Uno T, Sato H, Fujiwara A, et al. Beneficial effect of early intervention with garenoxacin for bacterial infection-induced acute exacerbation of bronchial asthma and chronic obstructive pulmonary disease. Int Arch Allergy Immunol 2019;178(4):355-362 [FREE Full text] [doi: 10.1159/000495761] [Medline: 30759444]
51. Rubio N, Parker RA, Drost EM, Pinnock H, Weir CJ, Hanley J, et al. Home monitoring of breathing rate in people with chronic obstructive pulmonary disease: observational study of feasibility, acceptability, and change after exacerbation. Int J Chron Obstruct Pulmon Dis 2017;12:1221-1231 [FREE Full text] [doi: 10.2147/COPD.S120706] [Medline: 28458534]
52. Pericleous P, van Staa TP. The use of wearable technology to monitor physical activity in patients with COPD: a literature review. Int J Chron Obstruct Pulmon Dis 2019;14:1317-1322 [FREE Full text] [doi: 10.2147/COPD.S193037] [Medline: 31354259]
53. Khan S, Ali S, Bermak A. Recent developments in printing flexible and wearable sensing electronics for health care applications. Sensors (Basel) 2019 Mar 11;19(5):1230 [FREE Full text] [doi: 10.3390/s19051230] [Medline: 30862062]
54. Wu R, Liaqat D, de Lara E, Son T, Rudzicz F, Alshaer H, et al. Feasibility of using a smartwatch to intensively monitor patients with chronic obstructive pulmonary disease: prospective cohort study. JMIR Mhealth Uhealth 2018 Jun 14;6(6):e10046 [FREE Full text] [doi: 10.2196/10046] [Medline: 29903700]
55. Colantonio S, Govoni L, Dellacà RL, Martinelli M, Salvetti O, Vitacca M. Decision making concepts for the remote, personalized evaluation of COPD patients' health status. Methods Inf Med 2015;54(3):240-247 [FREE Full text] [doi: 10.3414/ME13-02-0038] [Medline: 25502493]
56. Li B, Dong Q, Downen RS, Tran N, Jackson JH, Pillai D, et al. A wearable IoT aldehyde sensor for pediatric asthma research and managemen. Sens Actuators B Chem 2019 May 15;287:584-594 [FREE Full text] [doi: 10.1016/j.snb.2019.02.077] [Medline: 31938011]
57. Katwa U, Rivera E. Asthma management in the era of smart-medicine: devices, gadgets, apps and telemedicine. Indian J Pediatr 2018 Sep;85(9):757-762 [FREE Full text] [doi: 10.1007/s12098-018-2611-6] [Medline: 29524089]

58. Flannick J, Johansson S, Njølstad PR. Common and rare forms of diabetes mellitus: towards a continuum of diabetes subtypes. Nat Rev Endocrinol 2016 Jul;12(7):394-406 [FREE Full text] [doi: 10.1038/nrendo.2016.50] [Medline: 27080136]
59. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2013 Jan;36 Suppl 1:S67-S74 [FREE Full text] [doi: 10.2337/dc13-S067] [Medline: 23264425]
60. Tentolouris A, Eleftheriadou I, Tzeravini E, Athanasakis K, Anastasiou IA, Kyriopoulos J, et al. Management of diabetes mellitus and patients' attitude towards the disease: Data from a nationwide study in Greece. Diabetes Metab Syndr 2019;13(2):1159-1164 [FREE Full text] [doi: 10.1016/j.dsx.2019.01.030] [Medline: 31336459]
61. Vishwanathan V. Current and future perspective in the management of diabetes. J Indian Med Assoc 2002 Mar;100(3):181-183 [FREE Full text] [Medline: 12408280]
62. Wan EYF, Fung CSC, Jiao FF, Yu EYT, Chin WY, Fong DYT, et al. Five-year effectiveness of the multidisciplinary risk assessment and management programme-diabetes mellitus (RAMP-DM) on diabetes-related complications and health service uses-a population-based and propensity-matched cohort study. Diabetes Care 2018 Jan;41(1):49-59 [FREE Full text] [doi: 10.2337/dc17-0426] [Medline: 29138274]
63. Gao W, Brooks GA, Klonoff DC. Wearable physiological systems and technologies for metabolic monitoring. J Appl Physiol (1985) 2018 Mar 01;124(3):548-556. [doi: 10.1152/japplphysiol.00407.2017] [Medline: 28970200]
64. Sabu C, Henna T, Raphey V, Nivitha K, Pramod K. Advanced biosensors for glucose and insulin. Biosens Bioelectron 2019 Sep 15;141:111201. [doi: 10.1016/j.bios.2019.03.034] [Medline: 31302426]
65. Pandey R, Paidi SK, Valdez TA, Zhang C, Spegazzini N, Dasari RR, et al. Noninvasive monitoring of blood glucose with raman spectroscopy. Acc Chem Res 2017 Feb 21;50(2):264-272 [FREE Full text] [doi: 10.1021/acs.accounts.6b00472] [Medline: 28071894]
66. Pillalamarri SS, Huyett LM, Abdel-Malek A. Novel bluetooth-enabled tubeless insulin pump: a user experience design approach for a connected digital diabetes management platform. J Diabetes Sci Technol 2018 Nov;12(6):1132-1142. [doi: 10.1177/1932296818804802] [Medline: 30304951]
67. Elliott WJ. Systemic hypertension. Curr Probl Cardiol 2007 Apr;32(4):201-259 [FREE Full text] [doi: 10.1016/j.cpcardiol.2007.01.002] [Medline: 17398315]
68. Oparil S, Acelajado MC, Bakris GL, Berlowitz DR, Cífková R, Dominiczak AF, et al. Hypertension. Nat Rev Dis Primers 2018 Mar 22;4:18014 [FREE Full text] [doi: 10.1038/nrdp.2018.14] [Medline: 29565029]
69. Holt TR, Withington DE, Mitchell E. Which pressure to believe? A comparison of direct arterial with indirect blood pressure measurement techniques in the pediatric intensive care unit. Pediatric Critical Care Medicine 2011;12(6):e391-e394 [FREE Full text] [doi: 10.1097/pcc.0b013e3182230f43] [Medline: 21666539]
70. Cheng H, Sung S, Chen C, Yu W, Yang S, Guo C, et al. Guiding hypertension management using different blood pressure monitoring strategies (GYMNs study): comparison of three different blood pressure measurement methods: study protocol for a randomized controlled trial. Trials 2019 May 10;20(1):265 [FREE Full text] [doi: 10.1186/s13063-019-3366-8] [Medline: 31077229]
71. Zhou W, Jia T, Chen X. The status and prospect of wearable medical health monitoring. China Medical Devices 2017;32(06):1-5. [doi: 10.3969/j.issn.1674-1633.2017.06.001]
72. Qi L, Wen Y. Research and application progress of wearable blood pressure measurement device. China Medical Devices 2016;31(12):77-79. [doi: 10.3969/j.issn.1674-1633.2016.12.019]
73. Zheng Y, Poon CCY, Yan BP, Lau JYW. Pulse arrival time based cuff-less and 24-h wearable blood pressure monitoring and its diagnostic value in hypertension. J Med Syst 2016 Sep;40(9):195 [FREE Full text] [doi: 10.1007/s10916-016-0558-6] [Medline: 27447469]
74. Islam SMS, Cartledge S, Karmakar C, Rawstorn JC, Fraser SF, Chow C, et al. Validation and acceptability of a cuffless wrist-worn wearable blood pressure monitoring device among users and health care professionals: mixed methods study. JMIR Mhealth Uhealth 2019 Sep 14;7(10):e14706 [FREE Full text] [doi: 10.2196/14706] [Medline: 31628788]
75. Mc Ardle R, Del Din S, Galna B, Thomas A, Rochester L. Differentiating dementia disease subtypes with gait analysis: feasibility of wearable sensors? Gait Posture 2020 Feb;76:372-376 [FREE Full text] [doi: 10.1016/j.gaitpost.2019.12.028] [Medline: 31901765]
76. Lane CA, Hardy J, Schott JM. Alzheimer's disease. Eur J Neurol 2018 Jan;25(1):59-70 [FREE Full text] [doi: 10.1111/ene.13439] [Medline: 28872215]
77. Ricci M, Di Lazzaro G, Pisani A, Mercuri NB, Giannini F, Saggio G. Assessment of motor impairments in early untreated Parkinson's disease patients: the wearable electronics impact. IEEE J Biomed Health Inform 2020 Jan;24(1):120-130 [FREE Full text] [doi: 10.1109/JBHI.2019.2903627] [Medline: 30843855]
78. Monje MH, Foffani G, Obeso J, Sánchez-Ferro Á. new sensor and wearable technologies to aid in the diagnosis and treatment monitoring of Parkinson's disease. Annu Rev Biomed Eng 2019 Jun 04;21:111-143 [FREE Full text] [doi: 10.1146/annurev-bioeng-062117-121036] [Medline: 31167102]
79. Mannini A, Trojaniello D, Cereatti A, Sabatini A. A machine learning framework for gait classification using inertial sensors: application to older adults, post-stroke and huntington's disease patients. Sensors (Basel) 2016 Jan 21;16(1) [FREE Full text] [doi: 10.3390/s16010134] [Medline: 26805847]

80. Crupi R, Faetti T, Paradiso R. Preliminary evaluation of wearable wellness system for obstructive sleep apnea detection. Annu Int Conf IEEE Eng Med Biol Soc 2015;2015:4142-4145 [FREE Full text] [doi: 10.1109/EMBC.2015.7319306] [Medline: 26737206]
81. Surrel G, Aminifar A, Rincon F, Murali S, Atienza D. Online obstructive sleep apnea detection on medical wearable sensors. IEEE Trans Biomed Circuits Syst 2018 Aug;12(4):762-773 [FREE Full text] [doi: 10.1109/TBCAS.2018.2824659] [Medline: 29993894]
82. Odeneg T, Ebner C, Mörtl D, Keller H, Dirninger A, Stix G, et al. Indications for and outcome in patients with the wearable cardioverter-defibrillator in a nurse-based training programme: results of the Austrian WCD Registry. Eur J Cardiovasc Nurs 2019 Jan;18(1):75-83 [FREE Full text] [doi: 10.1177/1474515118790365] [Medline: 30064269]
83. Davies H. Wearable medical device improves survival for glioblastoma patients. Cancer 2017 Sep 15;123(18):3433 [FREE Full text] [doi: 10.1002/cncr.30963] [Medline: 28880388]
84. Gura V, Rivara M, Bieber S, Munshi R, Smith N, Linke L, et al. A wearable artificial kidney for patients with end-stage renal disease. JCI Insight 2016 Jun 02;1(8):74-89 [FREE Full text] [doi: 10.1172/jci.insight.86397] [Medline: 27398407]
85. Wu X, Liu R, Xu S, Yang C, Yang S, Shao Z, et al. Feasibility of mixed reality-based intraoperative three-dimensional image-guided navigation for atlanto-axial pedicle screw placement. Proc Inst Mech Eng H 2019 Dec;233(12):1310-1317 [FREE Full text] [doi: 10.1177/0954411919881255] [Medline: 31617820]
86. Wu X, Liu R, Yu J, Xu S, Yang C, Yang S, et al. Mixed reality technology launches in orthopedic surgery for comprehensive preoperative management of complicated cervical fractures. Surg Innov 2018 Aug;25(4):421-422 [FREE Full text] [doi: 10.1177/1553350618761758] [Medline: 30012077]
87. Wu X, Liu R, Yu J, Xu S, Yang C, Shao Z, et al. Mixed reality technology-assisted orthopedics surgery navigation. Surg Innov 2018 Jun 27;25(3):304-305 [FREE Full text] [doi: 10.1177/1553350618771413] [Medline: 29701134]
88. Popp WL, Schneider S, Bär J, Bösch P, Spengler CM, Gassert R, et al. Wearable sensors in ambulatory individuals with a spinal cord injury: from energy expenditure estimation to activity recommendations. Front Neurol 2019;10:1092 [FREE Full text] [doi: 10.3389/fneur.2019.01092] [Medline: 31736845]
89. Faria AL, Andrade A, Soares L, I Badia SB. Benefits of virtual reality based cognitive rehabilitation through simulated activities of daily living: a randomized controlled trial with stroke patients. J Neuroeng Rehabil 2016 Nov 02;13(1):96 [FREE Full text] [doi: 10.1186/s12984-016-0204-z] [Medline: 27806718]
90. Williamson B, Aplin T, de Jonge D, Goyne M. Tracking down a solution: exploring the acceptability and value of wearable GPS devices for older persons, individuals with a disability and their support persons. Disabil Rehabil Assist Technol 2017 Nov 18;12(8):822-831 [FREE Full text] [doi: 10.1080/17483107.2016.1272140] [Medline: 28100086]
91. Shi B, Chen X, Yue Z, Yin S, Weng Q, Zhang X, et al. Wearable ankle robots in poststroke rehabilitation of gait: a systematic review. Front Neurorobot 2019;13:63 [FREE Full text] [doi: 10.3389/fnbot.2019.00063] [Medline: 31456681]
92. van Genderen ME, Vlake J. [Virtual healthcare; use of virtual, augmented and mixed reality]. Ned Tijdschr Geneeskd 2018 Nov 26;162 [FREE Full text] [Medline: 30500123]
93. Likitlersuang J, Sumitro ER, Theventhiran P, Kalsi-Ryan S, Zariffa J. Views of individuals with spinal cord injury on the use of wearable cameras to monitor upper limb function in the home and community. J Spinal Cord Med 2017 Nov 24;40(6):706-714 [FREE Full text] [doi: 10.1080/10790268.2017.1349856] [Medline: 28738759]
94. Wang L, Liu J, Yang Y. The prevention and treatment of stroke still face huge challenges??brief report on stroke prevention and treatment in China, 2018. Chinese Circulation Journal 2019;34(02):105. [doi: 10.3969/j.issn.1000-3614.2019.02.001]
95. Parker VM, Wade DT, Langton Hewer R. Loss of arm function after stroke: measurement, frequency, and recovery. Int Rehabil Med 1986;8(2):69-73 [FREE Full text] [doi: 10.3109/03790798609166178] [Medline: 3804600]
96. Hsu W, Sugiarto T, Lin Y, Yang F, Lin Z, Sun C, et al. Multiple-wearable-sensor-based gait classification and analysis in patients with neurological disorders. Sensors (Basel) 2018 Oct 11;18(10):3397 [FREE Full text] [doi: 10.3390/s18103397] [Medline: 30314269]
97. Richard D W, Gerd L, Privender S, Annick T, Jurgen te V, Stefan W. Home stroke rehabilitation for the upper limbs. 2007 Aug 22 Presented at: 29th International Conference of the IEEE Engineering in Medicine and Biology Society; 2007; Lyon, France p. 4015-4018 URL: https://ieeexplore.ieee.org/document/4353214/citations#citations [doi: 10.1109/iembs.2007.4353214]
98. Maceira-Elvira P, Popa T, Schmid A, Hummel FC. Wearable technology in stroke rehabilitation: towards improved diagnosis and treatment of upper-limb motor impairment. J Neuroeng Rehabil 2019 Nov 19;16(1):142 [FREE Full text] [doi: 10.1186/s12984-019-0612-y] [Medline: 31744553]
99. Ciotti S, Battaglia E, Carbonaro N, Bicchi A, Tognetti A, Bianchi M. A synergy-based optimally designed sensing glove for functional grasp recognition. Sensors (Basel) 2016 Jun 02;16(6):811 [FREE Full text] [doi: 10.3390/s16060811] [Medline: 27271621]
100. Dimyan MA, Cohen LG. Neuroplasticity in the context of motor rehabilitation after stroke. Nat Rev Neurol 2011 Feb;7(2):76-85 [FREE Full text] [doi: 10.1038/nrneurol.2010.200] [Medline: 21243015]
101. Young AJ, Ferris DP. State of the art and future directions for lower limb robotic exoskeletons. IEEE Trans Neural Syst Rehabil Eng 2017 Feb;25(2):171-182 [FREE Full text] [doi: 10.1109/TNSRE.2016.2521160] [Medline: 26829794]

102. Lu Z, Chen X, Zhang X, Tong K, Zhou P. Real-time control of an exoskeleton hand robot with myoelectric pattern recognition. Int J Neural Syst 2017 Aug;27(5):1750009 [FREE Full text] [doi: 10.1142/S0129065717500095] [Medline: 27873553]
103. Fan H, Cao X, Yang X. Application and research of wearable devices in orthopedics. Chinese Journal of Bone and Joint Surgery 2018 Feb;11(2):154-160. [doi: 10.3969/j.issn.2095-9958.2018.02.016]
104. Mandell JC, Khurana B, Smith SE. Stress fractures of the foot and ankle, part 1: biomechanics of bone and principles of imaging and treatment. Skeletal Radiol 2017 Aug;46(8):1021-1029 [FREE Full text] [doi: 10.1007/s00256-017-2640-7] [Medline: 28374052]
105. Zhu Y, Nakamura M, Horiuchi T, Kohno H, Takahashi R, Terada H, et al. New wearable walking-type continuous passive motion device for postsurgery walking rehabilitation. Proc Inst Mech Eng H 2013 Jul;227(7):733-745 [FREE Full text] [doi: 10.1177/0954411913481557] [Medline: 23636753]
106. Lee S, Kim J, Baker L, Long A, Karavas N, Menard N, et al. Autonomous multi-joint soft exosuit with augmentation-power-based control parameter tuning reduces energy cost of loaded walking. J Neuroeng Rehabil 2018 Jul 13;15(1):66 [FREE Full text] [doi: 10.1186/s12984-018-0410-y] [Medline: 30001726]
107. Eshel I, Bowles AO, Ray MR. Rehabilitation of cognitive dysfunction following traumatic brain injury. Phys Med Rehabil Clin N Am 2019 Feb;30(1):189-206. [doi: 10.1016/j.pmr.2018.08.005] [Medline: 30470421]
108. Hwang J, Lee S. The effect of virtual reality program on the cognitive function and balance of the people with mild cognitive impairment. J Phys Ther Sci 2017 Aug;29(8):1283-1286 [FREE Full text] [doi: 10.1589/jpts.29.1283] [Medline: 28878448]
109. Chua SIL, Tan NC, Wong WT, Allen JC, Quah JHM, Malhotra R, et al. Virtual reality for screening of cognitive function in older persons: comparative study. J Med Internet Res 2019 Aug 01;21(8):e14821 [FREE Full text] [doi: 10.2196/14821] [Medline: 31373274]
110. Rogers JM, Duckworth J, Middleton S, Steenbergen B, Wilson PH. Elements virtual rehabilitation improves motor, cognitive, and functional outcomes in adult stroke: evidence from a randomized controlled pilot study. J Neuroeng Rehabil 2019 May 15;16(1):56 [FREE Full text] [doi: 10.1186/s12984-019-0531-y] [Medline: 31092252]
111. Chirico A, Lucidi F, De Laurentiis M, Milanese C, Napoli A, Giordano A. Virtual reality in health system: beyond entertainment. a mini-review on the efficacy of VR during cancer treatment. J Cell Physiol 2016 Feb;231(2):275-287 [FREE Full text] [doi: 10.1002/jcp.25117] [Medline: 26238976]
112. Faria AL, Andrade A, Soares L, I Badia SB. Benefits of virtual reality based cognitive rehabilitation through simulated activities of daily living: a randomized controlled trial with stroke patients. J Neuroeng Rehabil 2016 Nov 02;13(1):96 [FREE Full text] [doi: 10.1186/s12984-016-0204-z] [Medline: 27806718]
113. Wåhlin A, Fordell H, Ekman U, Lenfeldt N, Malm J. Rehabilitation in chronic spatial neglect strengthens resting-state connectivity. Acta Neurol Scand 2019 Mar;139(3):254-259 [FREE Full text] [doi: 10.1111/ane.13048] [Medline: 30427058]
114. Celik N, Rohrschneider K. [Electronic vision aids : New options for rehabilitation of the visually impaired]. Ophthalmologe 2018 Jul;115(7):553-558 [FREE Full text] [doi: 10.1007/s00347-017-0644-2] [Medline: 29322255]
115. Park Y, Lee S, Park J. Recent progress in wireless sensors for wearable electronics. Sensors (Basel) 2019 Oct 09;19(20):4353 [FREE Full text] [doi: 10.3390/s19204353] [Medline: 31600870]
116. Souza P, Arehart K, Schoof T, Anderson M, Strori D, Balmert L. Understanding variability in individual response to hearing aid signal processing in wearable hearing aids. Ear Hear 2019;40(6):1280-1292 [FREE Full text] [doi: 10.1097/AUD.0000000000000717] [Medline: 30998547]
117. Paterno L, Ibrahimi M, Gruppioni E, Menciassi A, Ricotti L. Sockets for limb prostheses: a review of existing technologies and open challenges. IEEE Trans Biomed Eng 2018 Sep;65(9):1996-2010 [FREE Full text] [doi: 10.1109/TBME.2017.2775100] [Medline: 29993506]
118. Segura Anaya LH, Alsadoon A, Costadopoulos N, Prasad PWC. Ethical implications of user perceptions of wearable devices. Sci Eng Ethics 2018 Feb;24(1):1-28 [FREE Full text] [doi: 10.1007/s11948-017-9872-8] [Medline: 28155094]
119. Lewy H. Wearable technologies - future challenges for implementation in healthcare services. Healthc Technol Lett 2015 Feb;2(1):2-5 [FREE Full text] [doi: 10.1049/htl.2014.0104] [Medline: 26609396]
120. Yetisen AK, Martinez-Hurtado JL, Ünal B, Khademhosseini A, Butt H. Wearables in medicine. Adv Mater 2018 Jun 11;30(33):e1706910 [FREE Full text] [doi: 10.1002/adma.201706910] [Medline: 29893068]
121. Dewsbury G, Clarke K, Rouncefield M, Sommerville I. Designing appropriate assistive technology for home users: Developing dependable networks. Presented at: Inclusive Design and Mobility Response in Indoor/Outdoor Public Buildings and Facilities: CIB Working Group W084 – Building Non-Handicapping Environments; October 21-22, 2002; Rome URL: https://tinyurl.com/y3foa8gs
122. Rodgers MM, Alon G, Pai VM, Conroy RS. Wearable technologies for active living and rehabilitation: current research challenges and future opportunities. J Rehabil Assist Technol Eng 2019;6:2055668319839607. [doi: 10.1177/2055668319839607] [Medline: 31245033]
123. Cilliers L. Wearable devices in healthcare: privacy and information security issues. Health Inf Manag 2020;49(2-3):150-156 [FREE Full text] [doi: 10.1177/1833358319851684] [Medline: 31146589]
124. Salim A, Lim S. Recent advances in noninvasive flexible and wearable wireless biosensors. Biosens Bioelectron 2019 Sep 15;141:111422. [doi: 10.1016/j.bios.2019.111422] [Medline: 31229794]

125. Ling Y, An T, Yap LW, Zhu B, Gong S, Cheng W. Disruptive, soft, wearable sensors. Adv Mater 2020 May 13;32(18):e1904664 [FREE Full text] [doi: 10.1002/adma.201904664] [Medline: 31721340]
126. Wolf JA, Moreau JF, Akilov O, Patton T, English JC, Ho J, et al. Diagnostic inaccuracy of smartphone applications for melanoma detection. JAMA Dermatol 2013 Apr;149(4):422-426 [FREE Full text] [doi: 10.1001/jamadermatol.2013.2382] [Medline: 23325302]

Abbreviations

VR: virtual reality
ECG: electrocardiogram
COPD: chronic obstructive pulmonary disease

Edited by G Eysenbach; submitted 26.03.20; peer-reviewed by D Cebo, KC Wong; comments to author 20.04.20; revised version received 22.09.20; accepted 24.09.20; published 09.11.20.

Please cite as:
Lu L, Zhang J, Xie Y, Gao F, Xu S, Wu X, Ye Z
Wearable Health Devices in Health Care: Narrative Systematic Review
JMIR Mhealth Uhealth 2020;8(11):e18907
URL: http://mhealth.jmir.org/2020/11/e18907/
doi:10.2196/18907
PMID:33164904

Review

The Use of Wearables in Clinical Trials During Cancer Treatment: Systematic Review

Ulrikke Lyng Beauchamp[1], BSc; Helle Pappot[1,2], MD, DMSc; Cecilie Holländer-Mieritz[1], MD

[1]Department of Oncology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

[2]Faculty of Health, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Cecilie Holländer-Mieritz, MD
Department of Oncology
Rigshospitalet
University of Copenhagen
Blegdamsvej 9
Copenhagen, 2100
Denmark
Phone: 45 35453545
Email: cecilie.hollaender-mieritz@regionh.dk

Abstract

Background: Interest in the use of wearables in medical care is increasing. Wearables can be used to monitor different variables, such as vital signs and physical activity. A crucial point for using wearables in oncology is if patients already under the burden of severe disease and oncological treatment can accept and adhere to the device. At present, there are no specific recommendations for the use of wearables in oncology, and little research has examined the purpose of using wearables in oncology.

Objective: The purpose of this review is to explore the use of wearables in clinical trials during cancer treatment, with a special focus on adherence.

Methods: PubMed and EMBASE databases were searched prior and up to October 3, 2019, with no limitation in the date of publication. The search strategy was aimed at studies using wearables for monitoring adult patients with cancer during active antineoplastic treatment. Studies were screened independently by 2 reviewers by title and abstract, selected for inclusion and exclusion, and the full-text was assessed for eligibility. Data on study design, type of wearable used, primary outcome, adherence, and device outcome were extracted. Results were presented descriptively.

Results: Our systematic search identified 1269 studies, of which 25 studies met our inclusion criteria. The types of cancer represented in the studies were breast (7/25), gastrointestinal (4/25), lung (4/25), and gynecologic (1/25); 9 studies had multiple types of cancer. Oncologic treatment was primarily chemotherapy (17/25). The study-type distribution was pilot/feasibility study (12/25), observational study (10/25), and randomized controlled trial (3/25). The median sample size was 40 patients (range 7-180). All studies used a wearable with an accelerometer. Adherence varied across studies, from 60%-100% for patients wearing the wearable/evaluable sensor data and 45%-94% for evaluable days, but was differently measured and reported. Of the 25 studies, the most frequent duration for planned monitoring with a wearable was 8-30 days (13/25). Topics for wearable outcomes were physical activity (19/25), circadian rhythm (8/25), sleep (6/25), and skin temperature (1/25). Patient-reported outcomes (PRO) were used in 17 studies; of the 17 PRO studies, only 9 studies reported correlations between the wearable outcome and the PRO.

Conclusions: We found that definitions of outcome measures and adherence varied across studies, and limited consensus among studies existed on which variables to monitor during treatment. Less heterogeneity, better consensus in terms of the use of wearables, and established standards for the definitions of wearable outcomes and adherence would improve comparisons of outcomes from studies using wearables. Adherence, and the definition of such, seems crucial to conclude on data from wearable studies in oncology. Additionally, research using advanced wearable devices and active use of the data are encouraged to further explore the potential of wearables in oncology during treatment. Particularly, randomized clinical studies are warranted to create consensus on when and how to implement in oncological practice.

(JMIR Mhealth Uhealth 2020;8(11):e22006) doi:10.2196/22006

KEYWORDS

cancer treatment; wearables; adherence; sensor technology

Introduction

Technology expansion over the past decade, along with the use of various sensors and electronic devices and the arrival of more advanced devices, has led to new possibilities [1-3]. The potential use of sensor technology in the health care setting is wide and covers aspects suitable for many purposes [4,5]. In this review, we focus on the use of wearables in clinical trials during cancer treatment.

A wearable is a device with a sensor that can collect health-related data remotely [6]. Depending on the device design, it can be worn in different ways, either on the wrist, upper arm, around the waist, fastened to the hip, or on another location of the body. Wearables can collect information on various biometric data points (eg, heart rate, respiratory rate, blood oxygen saturation, sleep pattern, or body temperature) [7]. This information can be used alone or in combination with other information [eg, patient-reported outcome (PRO) or other patient-generated health data] to evaluate or estimate a clinical outcome; hospitalization, adverse events, performance status, and physical activity can be used as a clinical outcome. The information can be used for home monitoring with feedback to the clinician or for self-management by the patient. The collection of data from the wearable can be either offline or in real time, depending on the type of feedback wanted for the user (eg, the patient or clinician).

In oncology, wearables may offer new vital information about patients, which can potentially lead to better management of cancer treatment [7]. This is important both in the aspect of precision care and economy, with cancer being the second leading cause of death globally [8]. Wearables make it possible to monitor patients at home in their own environment, compared to monitoring only at in-clinic visits [9]. Tracking data in the patient's own environment might allow the patient to continue their normal routines while their health data is being transmitted to a database or directly to the clinician. The possibilities are appealing in oncology from a "work smarter" approach and, more importantly, for the general improvements in the quality of life they may offer to patients with cancer of [10,11]. At present, there is little clinical evidence showing how wearables might improve the cancer pathway for patients with cancer during treatment [12].

To understand the potential use of wearables, it is relevant to capture what the wearable's objective outcome is, what effect or role it has on the clinical outcome, and what it can be used for in a medical setting [7]. Evaluating the patient's adherence to the wearable and defining valid data helps to ensure that new technologies are introduced into clinical practice with a focus on the patient's perspective. Limited consensus and guidelines exist for designing or reporting trials using wearables as part of the intervention—but this research area is getting increased attention [13,14]. One initiative is the Clinical Transformation Initiative (CTTI), which has issued recommendations regarding the appropriate use of mobile technology in clinical trials [13]. They have also initiated a database on feasibility studies in clinical trials [15]. The database is not limited to one specific disease group but only includes feasibility studies [15].

The purpose of this review is to explore the use of wearables in clinical trials during cancer treatment, with a focus on adherence and the setting.

Methods

Study Design and Search Strategy

Systematic searches were performed in PubMed and EMBASE. Both databases were searched prior and up to October 3, 2019, with no limitation in the date of publication. Searches consisted of cancer/neoplasm keywords and terms for wearable devices.

In PubMed, the search consisted of the medical-subject-heading (MeSH) terms "neoplasms," "medical oncology," "surgical oncology," and "wearable electronic devices," along with a combination of free-text words for the topics "oncology," "cancer," "wearable device," "accelerometer," and "actigraph." In EMBASE, the search included categorized terms for neoplasms and electronic monitoring devices such as "neoplasm," "patient monitoring," and "electronic device." Additional search terms "ambulatory monitoring" and "telemedicine" were added. The search was limited to articles published in English. The search strategy is shown in Multimedia Appendix 1. The review was registered at PROSPERO (International prospective register of systematic reviews) ID number CRD42020154386 before data extraction was initiated [16].

Criteria for Inclusion of Studies

Studies found with the selected search strategy were screened by title and abstract, which was performed independently by 2 reviewers who were blinded to each other's decisions. Cases of disagreement about whether to include or exclude a study were decided through a consensus decision. The included studies had their full text assessed for eligibility; disagreements were resolved by consensus, which was achieved in all cases.

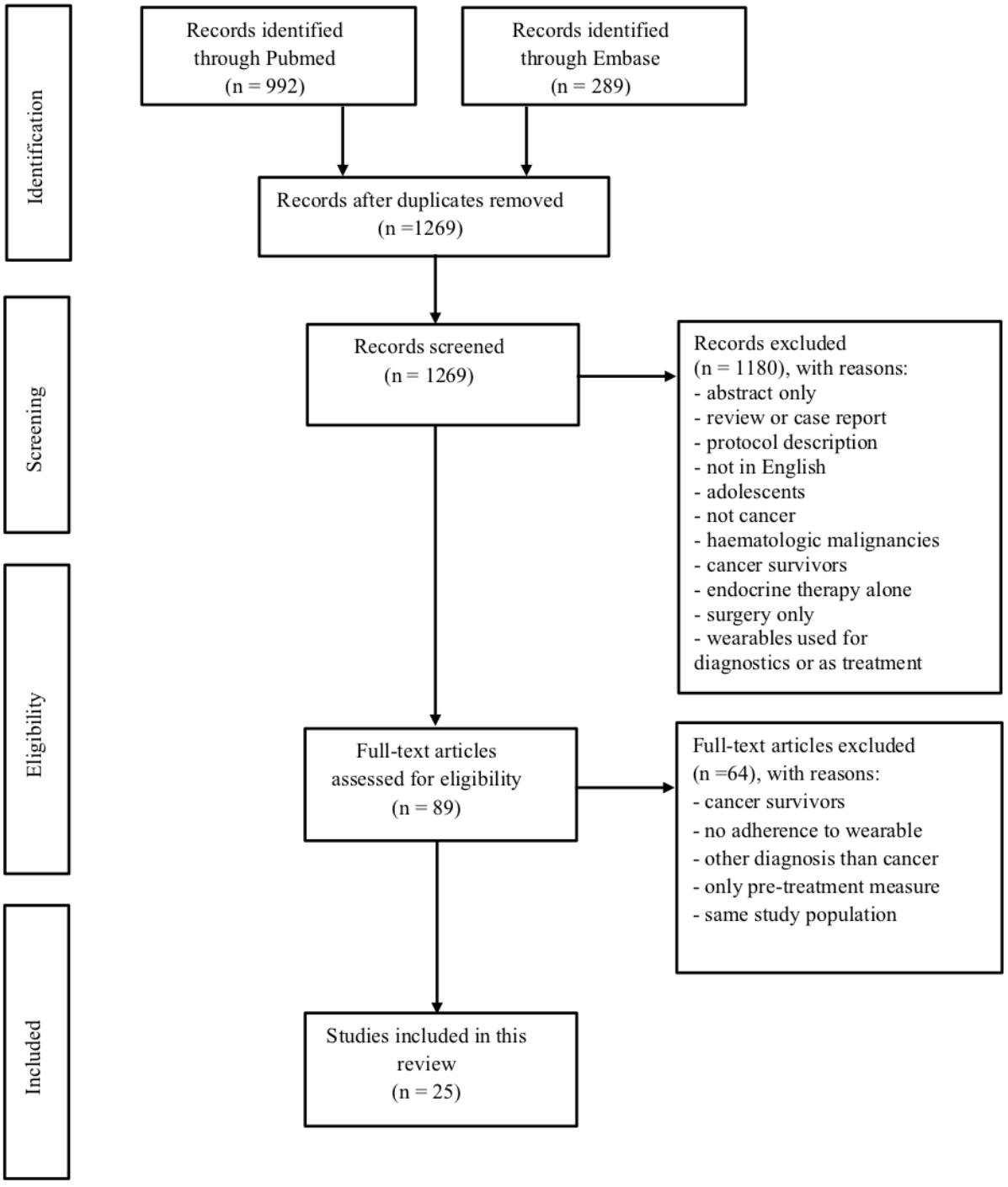
For a study to be included, it had to be written in English and be either a randomized controlled trial (RCT), observational study, or pilot study/ feasibility study. Patients had to be 18 years of age or older and diagnosed with a solid malignant tumor. It was mandatory that studies took place during active cancer treatment; treatment could be either radiation or antineoplastic treatment, such as chemotherapy or targeted therapy. Studies investigating all types of wearables were considered eligible if they had an objective measure. Studies had to include a description of adherence to the wearable to be eligible for inclusion.

Exclusion criteria were studies registered as protocol descriptions, study protocols, abstracts from conferences, editorials, letters, or case reports. Also excluded were studies in which patients were cancer survivors or had hematologic malignancies and were treated with endocrine therapy only or

surgery only. If the wearable devices were worn only pretreatment, the study used hearing aids as the wearable device, or wearables were used as treatment or for diagnostics, then these studies were also excluded.

The search is graphically presented according to the PRISMA flow diagram (Figure 1).

Figure 1. PRISMA flow diagram of the screening and selection of studies.



Data Extraction

The following study characteristics data were extracted: study title, author, year of publication, country, study design, number of patients included, and main objectives. Study population information that was extracted included age group, cancer type, treatment type, and intent of treatment. Study data regarding the adherence to the wearable were the type of wearable used (hardware, software), placement, device outcome, planned wear time, valid wear time, and adherence to the wearable. For the purposes of this review, the device outcome was defined as the objective measures used in the study (eg, step count); planned wear time was defined as the time period that patients were supposed to use the device; and valid wear time was defined as

the minimum wear time for data inclusion, extracted if available (eg, ≥10 hours/day). Adherence could either be the percentage of patients wearing the wearable for the period, the percentages of patients with evaluable sensor data, or the percentages of total evaluable days. Study outcomes were extracted and thematically grouped into wearable outcome, PRO, and clinical outcome. Wearable outcome was defined as circadian rhythm, physical activity (PA), skin temperature, and sleep. PRO topics were quality of life, PA, mental health, symptom registration, and others. Clinical outcomes included adverse events, performance status (PS), and hospitalization.

When reading through the full text of included studies, synonyms for "wearable" were registered and extracted from each study.

Outcomes and Analysis

The review was conducted to give a descriptive presentation of the use of wearables in clinical trials. The primary outcome was adherence to the wearable. The secondary outcomes were the study outcomes: the wearable outcome, the PRO, and the clinical outcome. The wearable outcome was subtracted from the device outcome by the authors, and thematically grouped as a wearable outcome. We also investigated whether the studies reported a relationship between the wearable outcome and the PRO. All data were presented descriptively.

Ethical Considerations

This review did not require national or institutional approval.

Results

The search strategies were performed in PubMed and EMBASE on October 3, 2019. A total of 1281 studies were identified through the searches. There were 12 duplicated records, which were excluded, leaving 1269 studies eligible for screening. Titles and abstracts were examined, which resulted in the exclusion of 1180 studies that did not meet our inclusion criteria. The remaining 89 studies were evaluated for eligibility. Through full-text access, 64 other irrelevant studies were excluded, leaving a total of 25 studies to be reviewed for the purpose of this review. This process of study screening and selection is illustrated in a PRISMA flow chart (Figure 1).

Study characteristics are presented in Table 1, which shows a wide heterogeneity with respect to primary cancer sites (range, 1-8 sites), treatment-type specifics (eg, whole-brain radiotherapy, to all treatments allowed), sample sizes (n=7-180), and age groups (22-94 years). In terms of cancer diagnosis, 13 of the 25 studies included patients with breast cancer, 7 having breast cancer exclusively. Gastrointestinal (GI) cancer was featured in 12 of the 25 studies, 4 of which studied patients with GI exclusively. Lung cancer was present among 10 of the 25 studies, 4 featuring this diagnosis exclusively. For 19 of the 25 studies, the patients were treated with curative intent. For most (17/25) of the included studies, treatment was chemotherapy. Of the 25 studies included, a total of 12 were pilot/feasibility studies, 10 were observational studies, and 3 were randomized controlled trials (Table 1).

In 5 of the 25 studies, the planned wear time was ≤7 days, 13 were between 8-30 days, 5 were between 31-90 days, and in 2 studies, the planned wear time was over 90 days (Table 2).

Table 1. Characteristics of included studies, n=25.

Study (year)	Country	Primary cancer site	Treatment type	Sample size, n	Age group (range or mean [SD])	Study type
Broderick et al (2019) [17]	United States	Mixed	Chemotherapy	42	24-72	Pilot / feasibility study
Champ et al (2018) [18]	United States	Breast	Radiotherapy	10	52-79	Pilot / feasibility study
Chevalier et al (2003) [19]	France	Gastrointestinal	Chemotherapy	10	43-73	Pilot / feasibility study
Dean et al (2013) [20]	United States	Lung	Chemotherapy	35	48-94	Observational study
Dreher et al (2019) [21]	United States	Breast	Chemotherapy	65	29-72	Observational study
Edbrooke et al (2019) [22]	Australia	Lung	Mixed	92	63 (12.3)	Randomized controlled trial
Gupta et al (2018) [23]	United States	Mixed	Systemic therapy	24	54 (12.5)	Pilot / feasibility study
Innominato et al (2016) [24]	United Kingdom	Mixed	Chemotherapy	31	35-91	Pilot / feasibility study
Li et al (2019) [25]	China	Breast	Adjuvant chemotherapy	180	22-74	Observational study
Low et al (2017) [11]	United States	Gastrointestinal	Chemotherapy	14	40-74	Pilot / feasibility study
Lowe et al (2014) [26]	Canada	Mixed	Radiotherapy (whole brain)	31	63.5 (10.4)	Observational study
Mouri et al (2018) [27]	Japan	Mixed	Chemotherapy	30	70-84	Pilot / feasibility study
Nyrop et al (2018) [28]	United States	Breast	Chemotherapy	100	24-64	Observational study
Ohri et al (2019) [29]	United States	Lung	Chemo-radiotherapy	50	38-90	Observational study
Ohri et al (2017) [30]	United States	Mixed	Chemo-radiotherapy	38	33-82	Pilot / feasibility study
Ortiz-Tudela et al (2014) [31]	France	Mixed	Chemotherapy	49	35-90	Observational study
Parker et al (2019) [32]	United States	Pancreas	Chemotherapy; chemo-radiotherapy	50	66 (8)	Observational study
Roche et al (2014) [33]	France	Gastrointestinal	Chemotherapy	16	51-89	Pilot / feasibility study
Roscoe et al (2002) [34]	United States	Breast	Chemotherapy +/- radiotherapy	102	34-79	Randomized controlled trial
Sarna et al (2001) [35]	United States	Mixed	Radiotherapy	7	48-74	Pilot / feasibility study
Savard et al (2009) [36]	United States	Breast	Chemotherapy	95	34-79	Observational study
Solk et al (2019) [37]	United States	Breast	Chemotherapy	67	31-71	Observational study
van der Meij et al (2012) [38]	The Netherlands	Lung	Chemo-radiotherapy	40	39-80	Randomized controlled trial
Vassbakk-Brovold et al (2016) [39]	Norway	Mixed	Chemotherapy	66	59 (11)	Pilot / feasibility study
Wright et al (2018) [40]	United States	Gynaecological	Chemotherapy	10	60 (11)	Pilot / feasibility study

Table 2. Description of wearables and adherence.

Planned wear time interval and study (year)	Hardware / software	Device outcome	Planned wear time / valid wear time	Adherence description
≤7 days				
Chevalier et al (2003) [19]	Actigraph, Ambulatory Monitoring Inc / Action 3.8	• Rest activity cycle (movements/period)	3 days / 3 days	• 100% (10/10) of the patients wore the device for the full period
Dean et al (2013) [20]	Motionlogger actigraph / Action 3	• Sleep efficiency (%) • Sleep (hours) • Wake after sleep onset (minutes)	7 days / not reported	• 86% (30/35) of the patients wore the device for the full period
Lowe et al (2014) [26]	activPAL[a] / not reported	• Position time (hours/day) • Energy expenditure (metabolic equivalent of task [MET] h/day) • Step count (steps/day)	7 days / not reported	• 77% (24/31) of the patients provided evaluable sensor data between 3 and 7 days
Roscoe et al (2002) [34]	Mini-Motionlogger Actigraph / Action 3	• Circadian consistency (I<O[b]) • Daytime activity level (minutes) • Sleep (%)	72 hours at 2 timepoints / not reported	• 89% (91/102) provided evaluable sensor data at second cycle of chemotherapy • 44% (45/102) provided evaluable sensor data at fourth cycle of chemotherapy
Vassbakk-Brovold et al 2016) [39]	SenseWear Armband Pro3 or SenseWear Armband Mini[a] / SenseWear version 6.1 for Pro3 and version 7.0 for Mini	• Physical activity (minutes/week) recorded in 1-minute epochs	5 days / ≥19.2 hrs, for ≥1 day	• 79 % (66/84) of the patients wore the device for the full period
8-30 days				
Edbrooke et al (2019) [22]	SenseWear accelerometer[a] / not reported	• Step count (steps/day) • Number of 10+ minutes step bouts/day • Duration of 10+ minutes bouts (minutes) • Cadence of 10+ minutes bouts (steps/min)	7 days at 3 timepoints / 8hrs/day, for ≥4 days	• 87% (80/92) of the patients provided evaluable sensor data at baseline • 71%(65/92) of the patients provided evaluable sensor data at 9 weeks • 60% (55/92) of the patients provided evaluable data at 6 months
Innominato et al (2016) [24]	Micro Motionlogger / Action 4	• Circadian rest-activity (I<O[b])	30 days / not reported	• Evaluable sensor data were available in 75 % of the total days (653/874)
Li et al (2019) [25]	GENEActiv Original / not reported	• Sleep efficiency (%) • Sleep duration (minutes) • Nighttime total wake time (minutes)	7 days at 3 timepoints / ≥5 days per timepoint	• 97% (175/180) of the patients provided evaluable sensor data at T2 • 76% (136/180) of the patients provided evaluable sensor data at T3
Low et al (2017) [11]	Fitbit Charge HR / not reported	• Step count (steps/day) • Floors climbed (n) • Sleep (minutes) • Awakenings (n) • Time in bed (minutes)	4 weeks / not reported	• Evaluable sensor data were available in 75 % of the total days (295/392 days)
Mouri et al (2018) [27]	Kenz Lifecorder - GS[a] / Lifelyzer - 05 coach	• Step count (steps/day) • Physical activity (minutes/day) (physical activity was rated ≥1.8 METs)	7 days at 3 timepoints / ≥5 hrs/day	• 93% (28/30) of the patients wore the device for the full period
Ohri et al (2019) [29]	Garmin Vivofit [a] / not reported	• Step count (steps/day)	Up to 3 weeks / not reported	• Evaluable sensor data were available in 94 % of the total days (741/791)

Planned wear time interval and study (year)	Hardware / software	Device outcome	Planned wear time / valid wear time	Adherence description
Ortiz-Tudela et al (2014) [31]	Mini-Motionlogger Actigraph / Action 4	• Rest-activity (I<O[b]) • Wrist accelerations (acc/minute)	10-14 days split into 4 periods of 3-4 days / not reported	• 86% (42/49) of the patients provided evaluable sensor data the full period
Roche et al (2014) [33]	Mini-Motionlogger and VitalSense / Action 4, version 1.10	• Rest-activity (I<O[b]) • Wrist accelerations (acc/minute) • Skin surface temperature (°C/minute)	12 days split into 3 periods of 4 days/ not reported	• 100% (16/16) of patients provided evaluable sensor data at baseline • 63% (10/16) of patients provided evaluable sensor data during therapy and after therapy administration
Sarna et al (2001) [35]	Actiwatch 2 / not reported	• Wrist movement (n/second) • Physical activity (15-minute intervals)	5 days at 2 timepoints/ ≥3 days per timepoint	• 100% (7/7) of the patients wore the device the full period
Savard et al (2009) [36]	Actillume / Action 3	• Circadian rhythm variables (calculated from orientation and movement)	72 hrs at 7 timepoints/ not reported	• 91% (86/95) of patients provided evaluable sensor data at baseline (first cycle of chemotherapy week 1: 80%, week 2: 73% and week 3: 76%; fourth cycle of chemotherapy week 1: 74%, week 2 63% and week 3: 68%)
Solk et al (2019) [37]	ActiGraph, model wGT3X-BT / ActiLife, version 6.13.3	• Activity data (1-minutes intervals)	10 days at 3 timepoints / ≥10 hrs/day	• 84% (63/75) of the patients provided evaluable sensor data for the full period
van der Meij et al (2012) [38]	PAM accelerometer, model AM101[a] / not reported	• Physical activity (index score, 3 points reflects about 10 min of walking)	7 days at 3 timepoints / ≥3 full days	• 65% (26/40) of the patients wore the device for the full period
Wright et al (2018) [40]	Fitbit Zip and Fitbit Charge 2 / Fitabase	• Step count (steps/day) • Heart rate	30 days / ≥4 days/week	• 90% (9/10) of the patients wore the devices for the full period
31-90 days				
Broderick et al (2019) [17]	Microsoft Band 2 / not reported	• Step count (steps/day) • Heart rate • Calories (calories/hour)	60 days / ≥6 hrs/day	• Evaluable sensor data were available 86 % of the days (only day 1-14 included)
Champ et al (2018) [18]	Misfit Shine[a] / not reported	• Step count (steps/day) • Calories (calories/day) • Walking distance (miles) • Sleep (hours)	10 weeks / not reported	• 90% (9/10) of the patients wore the device for the full period
Gupta A et al (2018) [23]	Fitbit Flex / not reported	• Step count (steps/day) • Physical activity (sedentary minutes/day) • Sleep (minutes)	12 weeks / ≥1 steps/day recorded	• 96% (23/24) wore the device for >50% of the period
Nyrop et al (2018) [28]	Fitbit Zip[a] / not reported	• Step count (steps/day)	6-12 weeks / ≥3 weeks	• 79% (100/127) of the patients provided evaluable sensor data
Ohri et al (2017) [30]	Garmin / not reported	• Step count (steps/day)	Up to 80 days / 80% of the days	• Evaluable sensor data were available 94 % of the days
>91 days				

Planned wear time interval and study (year)	Hardware / software	Device outcome	Planned wear time / valid wear time	Adherence description
Dreher et al (2019) [21]	Fitbit Charge HR or Fitbit Charge 2 / Fitabase	• Step count (steps/day) • Heart rate • Sleep data	Up to 270 days / ≥10 hrs/day	• Evaluable sensor data were available in 45% of the days across 9 months
Parker et al (2019) [32]	ActiGraph GT3X+[a] / ActiLife Software, Version 6	• Physical activity (minutes/week) (1-min epochs)	14 days at each therapy phase / ≥10 hrs/day, for ≥7 days per timepoint	• 88 % (44/50) of the patients provided evaluable sensor data

[a]Placement other than wrist (anterior mid-thigh, hip, triceps muscle waist, not reported).

[b]bI<O is computed as the percentage of activity epochs when in-bed (I), whose values are lower than the median level of activity when out-of-bed(O).

Adherence data, presenting how many patients were able to use or collect data from the wearable device, or how many evaluable days the wearable was worn, were collected. Adherence varied across studies, from 60%-100% and 45%-94%, respectively, but was differently measured and reported. Valid wear time was defined in 16 of the 25 studies. Different hardware and software were used. The most frequent placement of the wearable was the wrist.

In Table 3, study outcomes were grouped as wearable outcome, PRO, and clinical outcome, and their respective subtopics can be seen for each included study, showing that 1 study could have more than 1 topic assigned. Of the 25 studies included, 19 had the purpose of monitoring PA (wearable outcome) as an outcome. The second most frequent topic among wearable outcomes was circadian rhythm, monitored in 8 studies. With respect to PRO, 9 studies examined quality of life, 7 studied mental health, 7 studied physical activity, and 8 looked at symptoms. Clinical outcomes were few, 4 comprising performance status, 2 looking at adverse events, and 2 studying hospitalization (Table 3).

Table 3. Study outcomes

Cancer type and study (year)	Wearable outcomes				Patient-reported outcomes					Clinical outcomes			
	Circ. rhythm[a]	Phys. activity[b]	Skin temp.[c]	Sleep	Mental health	Phys. activity	QoL[d]	Symptoms	Other	Adverse events	Perf. status[e]	Hospitalization	Other
Breast													
Roscoe et al (2002) [34]	✓	✓		✓	✓				✓		✓		
Savard J et al (2009) [36]	✓			✓									
Champ et al (2017) [18]		✓		✓									
Li et al (2019) [25]	✓	✓											✓
Nyrop et al (2018) [28]		✓			✓	✓	✓	✓	✓	✓	✓		✓
Dreher et al (2019) [21]	✓	✓											
Solk et al (2019) [37]		✓			✓	✓		✓	✓				
Gastrointestinal													
Chevalier et al (2003) [19]	✓												
Roche et al (2014) [33]	✓		✓										
Low et al (2017) [11]		✓		✓				✓					
Parker et al (2019) [32]		✓				✓							✓
Gynecological													
Wright et al (2018) [40]		✓					✓	✓					
Lung													
van der Meij et al (2012) [38]		✓					✓						
Dean et al (2013) [20]				✓	✓		✓		✓				
Edbrooke et al (2019) [22]		✓			✓	✓	✓	✓					
Ohri et al (2019) [18]		✓										✓	✓
Mixed													
Sarna et al (2001) [35]		✓			✓			✓	✓				
Ortiz-Tudela et al (2014) [31]	✓									✓			
Lowe et al (2014) [26]		✓				✓	✓	✓					
Innominato et al (2016) [24]	✓							✓					

Cancer type and study (year)	Wearable outcomes				Patient-reported outcomes					Clinical outcomes			
	Circ. rhythm[a]	Phys. activity[b]	Skin temp.[c]	Sleep	Mental health	Phys. activity	QoL[d]	Symptoms	Other	Adverse events	Perf. status[e]	Hospitalization	Other
Vassbakk-Brovold et al (2016) [39]		✓				✓							✓
Ohri et al (2017) [30]		✓					✓					✓	✓
Gupta et al (2018) [23]		✓		✓	✓		✓		✓		✓		
Mouri et al (2018) [27]		✓					✓						
Broderick et al (2019) [17]		✓				✓			✓		✓		

[a]Circ. rhythm: Circadian rhythm.

[b]Phys. activity: Physical activity.

[c]Skin temp.: Skin temperature.

[d]QoL: Quality of life.

[e]Perf. status: Performance status.

Of the 17 PRO studies, only 9 studies reported correlations between the wearable outcome and the PRO (Table 4). It was primarily physical activity, which was compared with the PROs (7/9).

Synonyms for "wearable" were also collected for each study (data not shown) while reading through the full text, which reflected both the terms used to address the technology in general and the terms describing the actual device used in the study. The most commonly used term was "accelerometer," which was used in 12 studies; next was "actigraph," which was mentioned in 8 studies. The term "tracker" was used in 6 studies as a part of several terms, including "activity tracker," "wearable activity tracker," and "fitness tracker." The latter term was used similarly to "monitor," which was mentioned in 9 studies.

Table 4. Studies that reported relationships between wearable outcomes and patient-reported outcomes (PRO; n=9).

PRO	Wearable outcome			
	Circadian rhythm	Physical activity	Skin temperature	Sleep
Mental health	[34]	[28],[22], [23]	—[a]	[23]
Physical activity	—	[28], [22], [26], [39], [17]	—	—
Quality of life	—	[28], [22], [26], [23]	—	[23]
Symptoms	—	[28], [11], [22], [26]	—	[11]
Others	[34] (fatigue)	[28] (fatigue), [23] (fatigue), [17] (fatigue, sleep)	—	[40] (sleep), [23] (fatigue)

[a]No relationship reported.

Discussion

Summary of Findings

The use of wearable sensor devices has become a popular self-awareness gadget for many people today, especially when it comes to measuring physical activity [41]. In this review, we demonstrate the heterogeneous use of wearables during cancer treatment reported in research studies. In a search of the literature, 1269 studies were identified, of which 25 were included in our review. These studies represented different cancer types, with most focusing on mixed cancer types or breast cancer solely. Treatment given in the studies was primarily chemotherapy. Study types were pilot/feasibility, observational, and randomized controlled studies with sample sizes varying from 7 to 180 patients. All studies included in the review used a wearable with an accelerometer, but monitoring duration varied (3-270 days); and even though most studies (19/25) had physical activity as the wearable outcome, the device outcome for physical activity varied. With respect to adherence to using the wearable, a considerable variation was seen. Of the 17 PRO studies, 9 studies made comparisons between wearable outcome and PRO. In general, we noticed a broad variation in study designs, definitions, and outcomes within this field. This was also reflected in the number of synonyms for "wearable" used as terms to address the technology in general and as terms to describe the actual device used in a study.

Our Focus

Other studies have reviewed the use and effects of eHealth tools, such as those for patient self-reporting of medication management and use, and have concluded that more high-quality research is needed before standard implementation of such tools can occur [42]. In oncology, there is a growing urge to include patients in the management of their own illness, making wearable devices a valuable tool in cancer therapy. However, technical and clinical feasibility are essential aspects to explore [15], as the device measurements depend on patients using them. In this review, we have additionally focused on a somewhat understudied issue—adherence. We report adherence in relation to wear time and report how many patients were able to use or collect data with the wearable device, or how many evaluable days the wearable was worn. Adherence appears to have a wide range, and no data are available on missing wear time, which leaves us questioning if patients only wear wearables when they feel fit or when it is convenient. Thus, before designing large intervention studies using wearables, one must consider defining minimum wear time or conditions with mandatory wear time, since changes in these parameters might influence results [13,14,43].

Strengths and Limitations

To our knowledge, this is the first review of the use of wearables in clinical trials during cancer treatment. This review was limited to studies that included adherence as an issue, but was not restricted to specific types of wearables. However, we and others believe that the choice of wearable outcome is highly important. Determination of which variables to measure is crucial to ensure the purpose of the wearable when incorporated into patients' daily routines [1]. Outcomes must be of specific value regarding the individual patient's course of disease. We agree with other researchers that it is unnecessary to monitor many different kinds of variables if the outcome is not going to have a significant impact on the patient's treatment [41]. At present, most studies tend to choose PA as the wearable outcome, but this may be because knowledge from the field of fitness and fitness training has grown [44]. Other wearable outcomes might be more relevant in the cancer setting, but this remains to be studied. Further, the objective measure of PA across studies varies. Of 9 studies comparing the wearable outcome and PRO, only 2 were randomized controlled trials. Further investigations in this area is needed before conclusions can be drawn. We suggest that future research include measurements of relevant, well-defined outcomes and be based on guidelines within the field, where such exist [45,46].

Many different terms are being used to describe electronic devices for use in health care [47], which can make it difficult to get an overview of the field and to keep up with how far research has come. In general, the terms used to describe the technology differ, from being very specific (eg, "wearable activity tracker" or "physical activity monitor") to being broad, and covering several types of devices (eg, "telehealth," "mobile health," "wearable devices," and "remote patient monitoring") [7,48]. Consensus regarding the terms used could be helpful for both indexing studies and categorizing results by the specific types of devices used; such uniformity would improve research in the field and probably lead to more and improved knowledge.

The patient populations represented in this review mostly reflect breast and mixed cancer populations. Additionally, most studies are from the United States. It is questionable if results from such studies can be transferred to other diagnoses, countries, and cultural settings.

Implications for Future Research

As in this review, overall adherence to wear time or to wearable device interventions in general are difficult to compare. This is because almost every study has a different way of defining how many minutes or hours of wear time should count as a valid active day. By establishing standards for definitions of wear time, this could allow results across different patient populations to be compared more easily. This could also be solved by using a parameter other than step count as a measure of physical activity [41].

This review provides an overview of the frequently used wearables in oncology during therapy. Many of the wearables used have similar competencies, which might suggest the need to expand research into using more advanced wearables like smartwatches; according to Lu et al [2], very few clinical trials using smartwatches could be identified in 2016. Besides providing extended opportunities to measure variables, smartwatches also benefit from having a user-friendly display, which can make it feasible for patients to track their own activity status. Only 2 of the 25 reviewed studies used real-time feedback [24,40], and this feature might play an important role in motivating patients and possibly detecting worsening or severe symptoms earlier [11]. Technology allows plenty of different wearables to be applied in the oncology setting. However, it remains unknown if wearables can improve essential outcomes like overall survival or lead to other improvements in cancer treatment. Only future, well-designed research studies based on guidelines for this field with clinically relevant outcomes can help us decide when and where to apply these important tools.

This review provides an inventory for the status of wearables in clinical trials and can be used in addition to the CTTI studies database when designing new clinical trials with wearables [15].

Conclusion

This review provides an overview of the use of wearable devices in oncology care for patients with solid tumors receiving antineoplastic treatment. We extracted data from studies monitoring patients with cancer and presented these results specifically regarding adherence, the device outcomes, and the types of wearables used.

We found that definitions of outcome measures and adherence varied across studies, and limited consensus among studies existed on which variables to monitor during treatment.

Less heterogeneity and better consensus in terms of use and establishing standards for definitions of wearable outcomes and adherence would improve the comparisons of outcomes among studies using wearables. Adherence and consistent definitions are crucial for drawing conclusions from data from wearable studies in oncology. Additionally, research using advanced

wearable devices and active use of the data are encouraged to further explore the potential of wearables in oncology during treatment. Especially, randomized clinical studies are warranted to create consensus on when and how to implement in oncological practice.

Acknowledgments

CHM is supported by the Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[PDF File (Adobe PDF File), 62 KB - mhealth_v8i11e22006_app1.pdf]

References

1. Chung AE, Jensen RE, Basch EM. Leveraging Emerging Technologies and the. J Oncol Pract 2016 Oct;12(10):863-866 [FREE Full text] [doi: 10.1200/JOP.2016.015784] [Medline: 27624946]
2. Lu T, Fu C, Ma MH, Fang C, Turner AM. Healthcare Applications of Smart Watches. A Systematic Review. Appl Clin Inform 2016 Sep 14;7(3):850-869. [doi: 10.4338/ACI-2016-03-R-0042] [Medline: 27623763]
3. Pennell NA, Dicker AP, Tran C, Jim HSL, Schwartz DL, Stepanski EJ. mHealth: Mobile Technologies to Virtually Bring the Patient Into an Oncology Practice. Am Soc Clin Oncol Educ Book 2017;37:144-154. [doi: 10.1200/EDBK_176093] [Medline: 28561720]
4. Piwek L, Ellis DA, Andrews S, Joinson A. The Rise of Consumer Health Wearables: Promises and Barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
5. Jim HSL, Hoogland AI, Brownstein NC, Barata A, Dicker AP, Knoop H, et al. Innovations in research and clinical care using patient-generated health data. CA Cancer J Clin 2020 May;70(3):182-199 [FREE Full text] [doi: 10.3322/caac.21608] [Medline: 32311776]
6. Ravizza A, De Maria C, Di Pietro L, Sternini F, Audenino AL, Bignardi C. Comprehensive Review on Current and Future Regulatory Requirements on Wearable Sensors in Preclinical and Clinical Testing. Front Bioeng Biotechnol 2019;7:313 [FREE Full text] [doi: 10.3389/fbioe.2019.00313] [Medline: 31781554]
7. Dias D, Paulo Silva Cunha J. Wearable Health Devices—Vital Sign Monitoring, Systems and Technologies. Sensors 2018 Jul 25;18(8):2414. [doi: 10.3390/s18082414]
8. Cancer today 2018. World Health Organization: WHO. 2019 Dec 01. URL: https://gco.iarc.fr/today/data/factsheets/populations/900-world-fact-sheets.pdf [accessed 2019-12-01]
9. Beg MS, Gupta A, Stewart T, Rethorst CD. Promise of Wearable Physical Activity Monitors in Oncology Practice. JOP 2017 Feb;13(2):82-89. [doi: 10.1200/JOP.2016.016857]
10. Reilly CM, Bruner DW, Mitchell SA, Minasian LM, Basch E, Dueck AC, et al. A literature synthesis of symptom prevalence and severity in persons receiving active cancer treatment. Support Care Cancer 2013 Jun;21(6):1525-1550 [FREE Full text] [doi: 10.1007/s00520-012-1688-0] [Medline: 23314601]
11. Low CA, Dey AK, Ferreira D, Kamarck T, Sun W, Bae S, et al. Estimation of Symptom Severity During Chemotherapy From Passively Sensed Data: Exploratory Study. J Med Internet Res 2017 Dec 19;19(12):e420 [FREE Full text] [doi: 10.2196/jmir.9046] [Medline: 29258977]
12. Coughlin SS, Caplan LS, Stone R. Use of consumer wearable devices to promote physical activity among breast, prostate, and colorectal cancer survivors: a review of health intervention studies. J Cancer Surviv 2020 Jun;14(3):386-392. [doi: 10.1007/s11764-020-00855-1] [Medline: 31933148]
13. Developing Novel Endpoints Generated by Mobile Technologies for Use in Clinical Trials 2017. Clinical Trials Transformation Initiative. 2020 Jun 08. URL: https://www.ctti-clinicaltrials.org/files/novelendpoints-recs.pdf [accessed 2020-06-08]
14. Byrom B, Watson C, Doll H, Coons SJ, Eremenco S, Ballinger R, ePRO Consortium. Selection of and Evidentiary Considerations for Wearable Devices and Their Measurements for Use in Regulatory Decision Making: Recommendations from the ePRO Consortium. Value Health 2018 Jun;21(6):631-639 [FREE Full text] [doi: 10.1016/j.jval.2017.09.012] [Medline: 29909867]
15. Bakker JP, Goldsack JC, Clarke M, Coravos A, Geoghegan C, Godfrey A, et al. A systematic review of feasibility studies promoting the use of mobile technologies in clinical research. NPJ Digit Med 2019;2:47 [FREE Full text] [doi: 10.1038/s41746-019-0125-x] [Medline: 31304393]

16. 2020. International prospective register of systematic reviews, PROSPERO. 2020 May 20. URL: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=154386 [accessed 2020-05-20]
17. Broderick JE, May M, Schwartz JE, Li M, Mejia A, Nocera L, et al. Patient reported outcomes can improve performance status assessment: a pilot study. J Patient Rep Outcomes 2019 Jul 16;3(1):41 [FREE Full text] [doi: 10.1186/s41687-019-0136-z] [Medline: 31313047]
18. Champ CE, Ohri N, Klement RJ, Cantor M, Beriwal S, Glaser SM, et al. Assessing Changes in the Activity Levels of Breast Cancer Patients During Radiation Therapy. Clinical Breast Cancer 2018 Feb;18(1):e1-e6. [doi: 10.1016/j.clbc.2017.08.009]
19. Chevalier V, Mormont M, Curé H, Chollet P. Assessment of circadian rhythms by actimetry in healthy subjects and patients with advanced colorectal cancer. Oncol Rep 2003;10(3):733-737. [Medline: 12684651]
20. Dean GE, Redeker NS, Wang Y, Rogers AE, Dickerson SS, Steinbrenner LM, et al. Sleep, mood, and quality of life in patients receiving treatment for lung cancer. Oncol Nurs Forum 2013 Sep;40(5):441-451 [FREE Full text] [doi: 10.1188/13.ONF.441-451] [Medline: 23989018]
21. Dreher N, Hadeler EK, Hartman SJ, Wong EC, Acerbi I, Rugo HS, et al. Fitbit Usage in Patients With Breast Cancer Undergoing Chemotherapy. Clin Breast Cancer 2019 Dec;19(6):443-449.e1. [doi: 10.1016/j.clbc.2019.05.005] [Medline: 31285177]
22. Edbrooke L, Granger CL, Clark RA, Denehy L. Physical Activity Levels Are Low in Inoperable Lung Cancer: Exploratory Analyses from a Randomised Controlled Trial. J Clin Med 2019 Aug 23;8(9) [FREE Full text] [doi: 10.3390/jcm8091288] [Medline: 31450784]
23. Gupta A, Stewart T, Bhulani N, Dong Y, Rahimi Z, Crane K, et al. Feasibility of Wearable Physical Activity Monitors in Patients With Cancer. JCO Clin Cancer Inform 2018 Dec;2:1-10. [doi: 10.1200/CCI.17.00152] [Medline: 30652571]
24. Innominato PF, Komarzynski S, Mohammad-Djafari A, Arbaud A, Ulusakarya A, Bouchahda M, et al. Clinical Relevance of the First Domomedicine Platform Securing Multidrug Chronotherapy Delivery in Metastatic Cancer Patients at Home: The inCASA European Project. J Med Internet Res 2016 Nov 25;18(11):e305 [FREE Full text] [doi: 10.2196/jmir.6303] [Medline: 27888171]
25. Li W, Kwok CC, Chan DC, Ho AW, Ho C, Zhang J, et al. Disruption of sleep, sleep-wake activity rhythm, and nocturnal melatonin production in breast cancer patients undergoing adjuvant chemotherapy: prospective cohort study. Sleep Med 2019 Mar;55:14-21. [doi: 10.1016/j.sleep.2018.11.022] [Medline: 30743205]
26. Lowe SS, Danielson B, Beaumont C, Watanabe SM, Baracos VE, Courneya KS. Associations Between Objectively Measured Physical Activity and Quality of Life in Cancer Patients With Brain Metastases. Journal of Pain and Symptom Management 2014 Sep;48(3):322-332. [doi: 10.1016/j.jpainsymman.2013.10.012]
27. Mouri T, Naito T, Morikawa A, Tatematsu N, Miura S, Okayama T, et al. Promotion of Behavioral Change and the Impact on Quality of Life in Elderly Patients with Advanced Cancer: A Physical Activity Intervention of the Multimodal Nutrition and Exercise Treatment for Advanced Cancer Program. Asia Pac J Oncol Nurs 2018 Dec;5(4):383-390 [FREE Full text] [doi: 10.4103/apjon.apjon_21_18] [Medline: 30271820]
28. Nyrop KA, Deal AM, Choi SK, Wagoner CW, Lee JT, Wood A, et al. Measuring and understanding adherence in a home-based exercise intervention during chemotherapy for early breast cancer. Breast Cancer Res Treat 2017 Nov 9;168(1):43-55. [doi: 10.1007/s10549-017-4565-1]
29. Ohri N, Halmos B, Bodner WR, Cheng H, Guha C, Kalnicki S, et al. Daily Step Counts: A New Prognostic Factor in Locally Advanced Non-Small Cell Lung Cancer? Int J Radiat Oncol Biol Phys 2019 Nov 15;105(4):745-751. [doi: 10.1016/j.ijrobp.2019.07.055] [Medline: 31398385]
30. Ohri N, Kabarriti R, Bodner WR, Mehta KJ, Shankar V, Halmos B, et al. Continuous Activity Monitoring During Concurrent Chemoradiotherapy. International Journal of Radiation Oncology*Biology*Physics 2017 Apr;97(5):1061-1065. [doi: 10.1016/j.ijrobp.2016.12.030]
31. Ortiz-Tudela E, Iurisci I, Beau J, Karaboue A, Moreau T, Rol MA, et al. The circadian rest-activity rhythm, a potential safety pharmacology endpoint of cancer chemotherapy. Int J Cancer 2014 Jun 1;134(11):2717-2725 [FREE Full text] [doi: 10.1002/ijc.28587] [Medline: 24510611]
32. Parker NH, Ngo-Huang A, Lee RE, O'Connor DP, Basen-Engquist KM, Petzel MQB, et al. Physical activity and exercise during preoperative pancreatic cancer treatment. Support Care Cancer 2019 Jun;27(6):2275-2284. [doi: 10.1007/s00520-018-4493-6] [Medline: 30334105]
33. Roche VP, Mohamad-Djafari A, Innominato PF, Karaboué A, Gorbach A, Lévi FA. Thoracic surface temperature rhythms as circadian biomarkers for cancer chronotherapy. Chronobiol Int 2014 Apr;31(3):409-420 [FREE Full text] [doi: 10.3109/07420528.2013.864301] [Medline: 24397341]
34. Roscoe JA, Morrow GR, Hickok JT, Bushunow P, Matteson S, Rakita D, et al. Temporal interrelationships among fatigue, circadian rhythm and depression in breast cancer patients undergoing chemotherapy treatment. Support Care Cancer 2002 May;10(4):329-336. [doi: 10.1007/s00520-001-0317-0] [Medline: 12029433]
35. Sarna L, Conde F. Physical activity and fatigue during radiation therapy: a pilot study using actigraph monitors. Oncol Nurs Forum 2001 Jul;28(6):1043-1046. [Medline: 11475878]

36. Savard J, Liu L, Natarajan L, Rissling MB, Neikrug AB, He F, et al. Breast cancer patients have progressively impaired sleep-wake activity rhythms during chemotherapy. Sleep 2009 Sep;32(9):1155-1160 [FREE Full text] [doi: 10.1093/sleep/32.9.1155] [Medline: 19750920]
37. Solk P, Gavin K, Fanning J, Welch W, Lloyd G, Cottrell A, et al. Feasibility and acceptability of intensive longitudinal data collection of activity and patient-reported outcomes during chemotherapy for breast cancer. Qual Life Res 2019 Dec;28(12):3333-3346. [doi: 10.1007/s11136-019-02278-7] [Medline: 31493269]
38. van der Meij BS, Langius JAE, Spreeuwenberg MD, Slootmaker SM, Paul MA, Smit EF, et al. Oral nutritional supplements containing n-3 polyunsaturated fatty acids affect quality of life and functional status in lung cancer patients during multimodality treatment: an RCT. Eur J Clin Nutr 2012 Mar;66(3):399-404 [FREE Full text] [doi: 10.1038/ejcn.2011.214] [Medline: 22234041]
39. Vassbakk-Brovold K, Kersten C, Fegran L, Mjåland O, Mjåland S, Seiler S, et al. Cancer patients participating in a lifestyle intervention during chemotherapy greatly over-report their physical activity level: a validation study. BMC Sports Sci Med Rehabil 2016;8:10 [FREE Full text] [doi: 10.1186/s13102-016-0035-z] [Medline: 27099757]
40. Wright AA, Raman N, Staples P, Schonholz S, Cronin A, Carlson K, et al. The HOPE Pilot Study: Harnessing Patient-Reported Outcomes and Biometric Data to Enhance Cancer Care. JCO Clin Cancer Inform 2018 Dec;2:1-12. [doi: 10.1200/CCI.17.00149] [Medline: 30652585]
41. Wood WA, Basch E. From Intuition to Execution: Realizing the Potential of Wearables in Oncology. J Oncol Pract 2017 Feb;13(2):90-92. [doi: 10.1200/JOP.2016.020370] [Medline: 28972831]
42. Lancaster K, Abuzour A, Khaira M. The Use and Effects of Electronic Health Tools for Patient Self-Monitoring and Reporting of Outcomes Following Medication Useystematic Review. Journal of medical Internet research Dec 18;. PubMed 2018;20(12):e294. [doi: 10.2196/jmir.9284] [Medline: 30563822]
43. Cox SM, Lane A, Volchenboum SL. Use of Wearable, Mobile, and Sensor Technology in Cancer Clinical Trials. JCO Clin Cancer Inform 2018 Dec;2:1-11. [doi: 10.1200/CCI.17.00147] [Medline: 30652590]
44. Brickwood K, Watson G, O'Brien J, Williams AD. Consumer-Based Wearable Activity Trackers Increase Physical Activity Participation: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2019 Apr 12;7(4):e11819 [FREE Full text] [doi: 10.2196/11819] [Medline: 30977740]
45. Goldsack JC, Coravos A, Bakker JP, Bent B, Dowling AV, Fitzer-Attas C, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). NPJ Digit Med 2020;3:55 [FREE Full text] [doi: 10.1038/s41746-020-0260-4] [Medline: 32337371]
46. Holländer-Mieritz C, Johansen C, Pappot H. eHealth-mind the gap. Acta Oncol 2020 Aug;59(8):877-878. [doi: 10.1080/0284186X.2020.1794037] [Medline: 32687421]
47. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote Patient Monitoring via Non-Invasive Digital Technologies: A Systematic Review. Telemed J E Health 2017 Jan;23(1):3-17 [FREE Full text] [doi: 10.1089/tmj.2016.0051] [Medline: 27116181]
48. Weinstein RS, Lopez AM, Joseph BA, Erps KA, Holcomb M, Barker GP, et al. Telemedicine, telehealth, and mobile health applications that work: opportunities and barriers. Am J Med 2014 Mar;127(3):183-187. [doi: 10.1016/j.amjmed.2013.09.032] [Medline: 24384059]

Abbreviations

CTTI: Clinical Transformation Initiative
GI: gastrointestinal
GYN: gynecological
PA: physical activity
PRO: patient-reported outcome
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 02.07.20; peer-reviewed by W Wood; comments to author 10.07.20; revised version received 04.09.20; accepted 07.09.20; published 11.11.20.

Please cite as:
Beauchamp UL, Pappot H, Holländer-Mieritz C
The Use of Wearables in Clinical Trials During Cancer Treatment: Systematic Review
JMIR Mhealth Uhealth 2020;8(11):e22006
URL: http://mhealth.jmir.org/2020/11/e22006/
doi:10.2196/22006
PMID:33174852

Corrigenda and Addenda

Correction: Excessive Smartphone Use and Self-Esteem Among Adults With Internet Gaming Disorder: Quantitative Survey Study

Hyunmin Kim[1,2], PhD; In Young Choi[1], PhD; Dai-Jin Kim[3], MD, PhD

[1]Department of Medical Informatics, College of Medicine, The Catholic University of Korea, Seoul St. Mary's Hospital, Seoul, Republic of Korea
[2]Division of Health Systems Management and Policy, School of Public Health, The University of Memphis, Memphis, TN, United States
[3]Department of Psychiatry, College of Medicine, The Catholic University of Korea, Seoul St. Mary's Hospital, Seoul, Republic of Korea

Corresponding Author:
In Young Choi, PhD
Department of Medical Informatics
College of Medicine
The Catholic University of Korea, Seoul St. Mary's Hospital
222 Banpo-daero, Seocho-gu
Seoul
Republic of Korea
Phone: 82 2258 7870
Email: iychoi@catholic.ac.kr

Related Article:

Correction of: https://mhealth.jmir.org/2020/9/e18505/

(JMIR Mhealth Uhealth 2020;8(11):e24869) doi:10.2196/24869

In "Excessive Smartphone Use and Self-Esteem Among Adults With Internet Gaming Disorder: Quantitative Survey Study" (JMIR Mhealth Uhealth 2020;8(9):e18505), the authors noted one error in the Acknowledgments section.

In the originally published manuscript, the final sentence of the Acknowledgments section read "This work was supported by the National Research Foundation of Korea (NRF) (grant no. 2014M3C7A1062893)." The has been changed to: "This work was supported by the National Research Foundation of Korea (NRF) (grant nos. 2014M3C7A1062893 and NRF-2019R1A5A2027588)."

The correction will appear in the online version of the paper on the JMIR Publications website on November 3, 2020, 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 08.10.20; this is a non–peer-reviewed article;accepted 20.10.20; published 03.11.20.

Please cite as:
Kim H, Choi IY, Kim DJ
Correction: Excessive Smartphone Use and Self-Esteem Among Adults With Internet Gaming Disorder: Quantitative Survey Study
JMIR Mhealth Uhealth 2020;8(11):e24869
URL: http://mhealth.jmir.org/2020/11/e24869/
doi:10.2196/24869
PMID:33141680

Review

Citizen-Centered Mobile Health Apps Collecting Individual-Level Spatial Data for Infectious Disease Management: Scoping Review

Felix Nikolaus Wirth[1,2], MSc; Marco Johns[1,2], MSc; Thierry Meurers[1,2], MSc; Fabian Prasser[1,2], PhD

[1]Berlin Institute of Health, Berlin, Germany

[2]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

Corresponding Author:
Felix Nikolaus Wirth, MSc
Berlin Institute of Health
Charitéplatz 1
Berlin, 10117
Germany
Phone: 49 15773326360
Email: felix-nikolaus.wirth@charite.de

Abstract

Background: The novel coronavirus SARS-CoV-2 rapidly spread around the world, causing the disease COVID-19. To contain the virus, much hope is placed on participatory surveillance using mobile apps, such as automated digital contact tracing, but broad adoption is an important prerequisite for associated interventions to be effective. Data protection aspects are a critical factor for adoption, and privacy risks of solutions developed often need to be balanced against their functionalities. This is reflected by an intensive discussion in the public and the scientific community about privacy-preserving approaches.

Objective: Our aim is to inform the current discussions and to support the development of solutions providing an optimal balance between privacy protection and pandemic control. To this end, we present a systematic analysis of existing literature on citizen-centered surveillance solutions collecting individual-level spatial data. Our main hypothesis is that there are dependencies between the following dimensions: the use cases supported, the technology used to collect spatial data, the specific diseases focused on, and data protection measures implemented.

Methods: We searched PubMed and IEEE Xplore with a search string combining terms from the area of infectious disease management with terms describing spatial surveillance technologies to identify studies published between 2010 and 2020. After a two-step eligibility assessment process, 27 articles were selected for the final analysis. We collected data on the four dimensions described as well as metadata, which we then analyzed by calculating univariate and bivariate frequency distributions.

Results: We identified four different use cases, which focused on individual surveillance and public health (most common: digital contact tracing). We found that the solutions described were highly specialized, with 89% (24/27) of the articles covering one use case only. Moreover, we identified eight different technologies used for collecting spatial data (most common: GPS receivers) and five different diseases covered (most common: COVID-19). Finally, we also identified six different data protection measures (most common: pseudonymization). As hypothesized, we identified relationships between the dimensions. We found that for highly infectious diseases such as COVID-19 the most common use case was contact tracing, typically based on Bluetooth technology. For managing vector-borne diseases, use cases require absolute positions, which are typically measured using GPS. Absolute spatial locations are also important for further use cases relevant to the management of other infectious diseases.

Conclusions: We see a large potential for future solutions supporting multiple use cases by combining different technologies (eg, Bluetooth and GPS). For this to be successful, however, adequate privacy-protection measures must be implemented. Technologies currently used in this context can probably not offer enough protection. We, therefore, recommend that future solutions should consider the use of modern privacy-enhancing techniques (eg, from the area of secure multiparty computing and differential privacy).

(JMIR Mhealth Uhealth 2020;8(11):e22594) doi:10.2196/22594

KEYWORDS

pandemic; epidemic; infectious disease management; mobile apps; automated digital contact tracing; mobility tracking; outbreak detection; location-based risk assessment; public health; informatics; app; infectious disease; COVID-19; review

Introduction

Background

In December 2019, a novel coronavirus (SARS-CoV-2) appeared and rapidly spread around the world. COVID-19, the disease associated with the virus, can cause severe respiratory illness, is highly transmissible among humans [1], and is often difficult to detect due to asymptomatic courses [2]. The outbreak was declared a Public Health Emergency of International Concern and has implications for global health and economic development alike.

To contain the pandemic, nonpharmaceutical interventions have been implemented on a broad scale as a rapid response. They aim at slowing down or interrupting the infection process by breaking the chain of infections [3]. Examples include individual measures such as social distancing, the wearing of masks, and self-isolation [4] as well as public measures including travel restrictions [5], closure of public institutions and businesses [6], quarantines, and curfews [7]. Many of these measures, however, come with drastic socioeconomic consequences [8]. For example, it has been estimated that the reduction in economic activity resulting from the lockdown will have an impact on the Euro area that is three to four times larger than the impact of the global financial crisis of 2007 and 2008 [9]. Further, it is expected that the pandemic and subsequent lockdowns will lead to mental health problems [10] and excess mortality indirectly related to COVID-19 [11]. Hence, there is considerable pressure to move toward policies that will allow a return of economic and social life to the former state while effectively containing the spread of SARS-CoV-2. This can probably best be achieved by replacing large-scale interventions with local or even individual interventions based on testing, tracing, and targeted quarantine [12,13]. This, however, requires effective means of disease surveillance.

For example, contact tracing is traditionally the manual process of identifying past contacts of a person with an infectious disease [14]. However, manual contact tracing is time and labor intensive, which is a challenge considering the scale of the COVID-19 pandemic. Thus, many believe that citizen-centered participatory digital tools are required [15]. Most notably, automated digital contact tracing via mobile apps (henceforth simply called contact tracing) is discussed in public media and scientific literature alike [16-20], and according infrastructures have been implemented in several countries. At the time of writing, the COVID Tracing Tracker Project lists more than 40 different apps [21].

An important aspect of automated contact tracing is its reliance on citizen-centered mobile apps collecting individual-level spatial data. The broad adoption of such apps is considered a major prerequisite for associated interventions to be effective [22], and it has been argued that data security and privacy aspects are a critical factor for adoption [22,23]. This is reflected by an intense discussion in the public and the scientific community about privacy-preserving solutions to the contact tracing problem [24,25] and the multitude of established solutions and projects [26-30]. However, often there is a trade-off between privacy risks associated with and the functionalities provided by solutions developed. The German contact tracing app, for example, focuses solely on contacts and builds upon an infrastructure with few central services [31]. As a result, it only supports contact tracing and does not offer any additional functionalities for disease surveillance.

There are, however, several additional use cases that can be implemented with location-based mobile health apps that could help with managing the current and future pandemics. A few unsystematic overviews of technologies currently implemented to manage the COVID-19 pandemic have been published [23,32,33]. For example, Boules and Geraghty [16] have listed several projects: HealthMap, a system that aggregates and maps informal information sources such as online social networks to create risk maps; WorldPop and EpiRis, which use location-based services and other sources of data to model human mobility to predict the spread of COVID-19; and the close contact detector, a solution that uses data on the movement of people to identify individuals with close contact to people who are infected. Other authors suggested the use of digital thermometers, smart watches, or other mobile health devices to remotely monitor vital signs and potential symptoms of at-risk individuals [32,33]. In addition, there are reviews focusing on solutions for health care professionals [34,35]. What is lacking, however, is a structured analysis of important aspects relevant to the technical properties and functionalities of current and future citizen-centered solutions.

Objectives

The objectives of this work are to inform the current discussions and to support the development of participatory disease surveillance technologies, providing an optimal balance between privacy protection and pandemic control. In our opinion, these objectives are best met by a systematic analysis of existing literature. Our main hypothesis is that there are dependencies between the specific diseases focused on, use cases supported, technology used, and data protection measures implemented. Furthermore, we hypothesize that these dependencies influence the possible design space of solutions. In particular, we focus on the following dimensions:

- *Use case dimension:* Many of the current discussions are focused on technical aspects and their consequences in terms of privacy risks. When considering use cases, the discussion seems to be focused only on contact tracing. Hence, we aim to obtain a broader view of the existing solutions' functionalities to facilitate the evaluation of novel concepts and to guide the development of future solutions.
- *Technology dimension:* Many recent developments are based on Bluetooth Low Energy (BLE), which is a relatively new technology, and GPS receivers. However, the choice of technology can also impact the functionalities that can

be implemented. To potentially broaden the design space for future tools, our objective is to get a broad overview of the technologies that have been used in prior projects and to analyze potential relationships to the use cases that have been implemented.
- *Disease dimension:* There are specific properties of SARS-CoV-2 that have contributed to the characteristic nature of the current pandemic. To determine whether existing solutions can serve as a blueprint for future work, we aim to investigate use cases and technologies in relationship to disease properties such as the path of transmission, infectiousness, and disease-associated symptoms.
- *Data protection dimension:* Finally, privacy protection is seen as an essential factor in current discussions on apps related to COVID-19. To derive a complete picture of data protection measures that can be implemented in this context, we aim to compile an overview of measures implemented in prior projects.

Methods

We performed a scoping review, as this type of review is best suited to map research activities in a broad and heterogeneous field [36]. Where applicable, we followed the guidelines of the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for scoping reviews) [37]. No review protocol was registered for this study. From the objectives, we derived the following inclusion criteria for articles:

- *Infectious disease management:* Solutions described must be focused on managing epidemics or pandemics of a pathogen, which is communicable directly from human-to-human, such as COVID-19, or transmitted by a vector, such as Dengue.
- *Mobile health apps:* Solutions proposed must be centered around a mobile device or app, which is used for data collection. Articles may describe a concept or an actual implementation. For instance, articles describing methods for modeling population-level disease spread but not covering methods of data collection were not included.
- *Based on individual-level spatial data:* Solutions described must be enabled by collecting individual-level spatial data (individual persons or small groups such as families), either absolute (eg, location determined via GPS) or relative (eg, contacts determined using Bluetooth), either dynamically (eg, regular updates to track movements) or statically (eg, manual entry of living address).
- *Citizen-centered solution:* Data must be collected either automatically or manually by regular citizens using their own devices. In contrast, solutions focused on data collection by professionals within a health care or surveillance context (such as apps for professional contact tracers) were excluded.

An article needed to fulfill all these criteria to be included. Furthermore, all papers were required to be written in English, published between the years 2010 and 2020 (final search performed on June 19, 2020), and scientific peer-reviewed papers containing original work.

We searched both PubMed and IEEE Xplore, as the topic of this review is placed at the intersection of medicine and technology. First, we specified a set of keywords describing the general context. At least one of these keywords needed to be contained in the title of articles checked for eligibility: *epidemic*, *pandemic*, *contact tracing*, *proximity tracing*, *surveillance*, or *infectious*. Second, we specified a set of keywords covering the complete spectrum of relevant technology. At least one of these keywords needed to be contained in either the title or the abstract of the articles: *mobile*, *wearables*, *smartphone*, *app*, *cellular network*, *NFC* (near-field communication), *barcode*, *QR* (Quick Response) *code*, *GPS*, *Wi-Fi*, *Bluetooth*, *RFID* (radio-frequency identification), *or magnetometer*. Search results were exported as comma-separated values files, harmonized, and imported into a consolidated spreadsheet. In total, we found 1133 articles, of which 646 were identified via PubMed and 487 via IEEE Xplore.

Figure 1 shows an overview of the selection process. The articles were selected in two steps. In each step, articles were randomly assigned to two authors for assessment, and cases of disagreement were discussed among all authors for consent. In a first screening step, the title and abstract were evaluated regarding the eligibility criteria. Articles were kept if the title or abstract did not provide for a decision on exclusion or inclusion. In total, 99 articles were included after this screening step. Next, the full texts of the selected articles were compared against the eligibility criteria. After this analysis, 27 articles were regarded as finally relevant and were included into the data extraction and analysis process. In this process, we extracted the year and free text for the remaining variables and mutually created categories from the text extracted. The data is available in Multimedia Appendix 1.

Table 1 illustrates the data items collected from the final selection of articles to answer the research questions previously outlined.

Figure 1. Overview of the selection process.

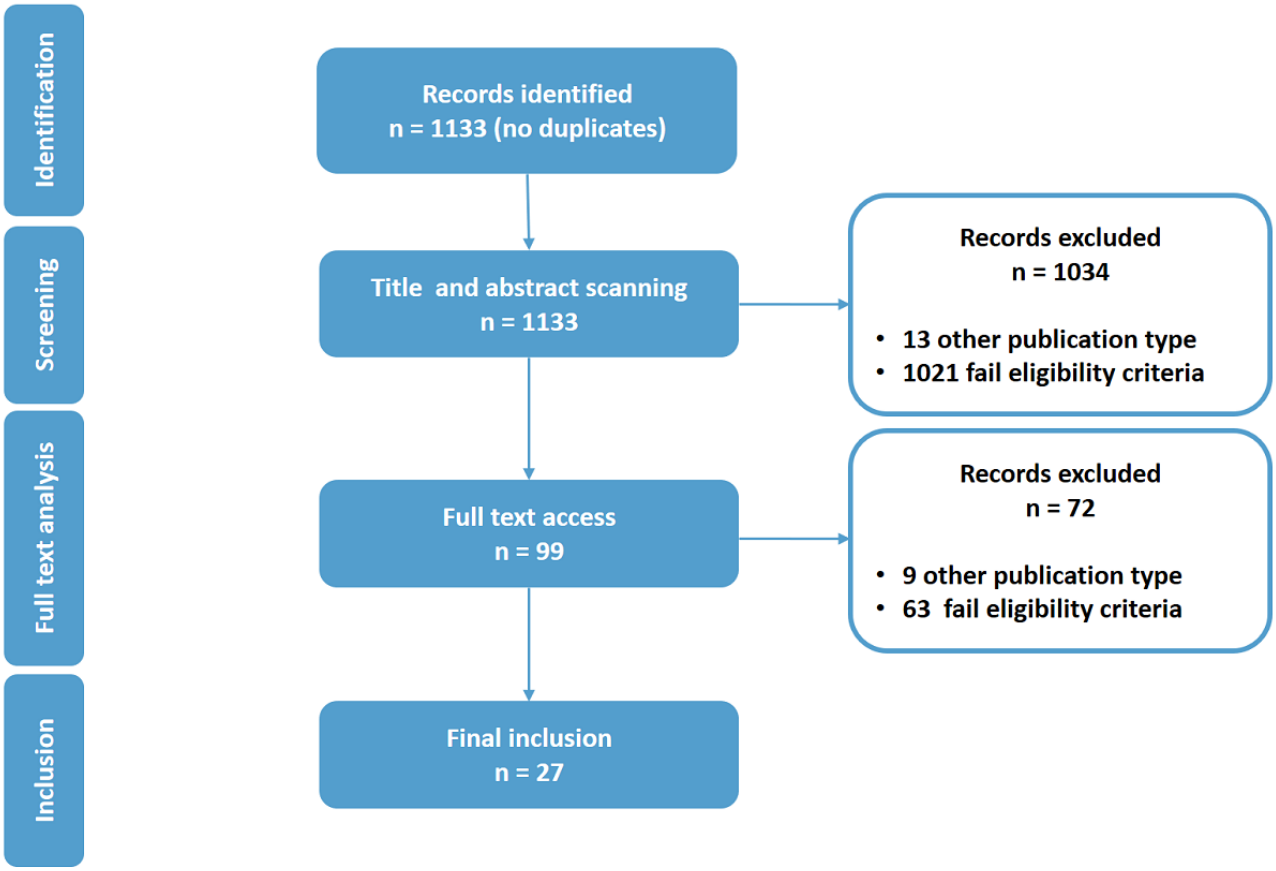


Table 1. Data items collected.

Variable	Examples	Definition
Year	2010, 2018,...	The year the article describing the solution was published
Use cases	Contact tracing, outbreak detection,...	The infectious disease management processes addressed by the solution
Sensor technologies	GPS, Bluetooth,...	The technology used for spatial tracking
Disease	COVID-19, influenza-like illness,...	The disease the solution focused on
Data protection	Geospatial aggregation, pseudonymization,...	Measures applied to protect the privacy of the users (technical and organizational)

The data collected corresponds with the dimensions investigated. First, we collected the year of publication to analyze developments over time. Second, in accordance with the first dimension, data was collected on the use case or cases supported by the apps. Third, in accordance with the second dimension, the sensor technology used to obtain spatial data was recorded. Fourth, in accordance with the third dimension, we collected data on the diseases that the solutions aimed to manage. Finally, in accordance with the fourth dimension, we recorded the data protection mechanisms used.

During data collection, the spreadsheet document was extended to work as a chartering form containing all data items. In accordance with our research questions, data was then analyzed by calculating frequency distributions of variables' values and by analyzing the relationships between variables using heat maps, visualizing the combined frequencies of their values.

Results

Overview

Table 2 shows the final selection of articles together with the collected data.

Table 2. Selected articles and data collected.

Author	Year	Use case	Disease	Sensor technology	Data protection measures
Abbas and Michael [38]	2020	Contact tracing	COVID-19	Bluetooth	Data minimization, pseudonymization
Abeler et al [24]	2020	Contact tracing	COVID-19	Bluetooth	Pseudonymization
Ackley et al [39]	2020	Outbreak detection	ILI[a]	GPS, IP[b] address geolocation	Geospatial aggregation, pseudonymization, temporal aggregation
Barrat et al [40]	2014	Contact tracing	Generic	Bluetooth	None
Chan et al [41]	2020	Outbreak detection	COVID-19	Manual entry	Consent, data minimization, pseudonymization
Farrahi et al [42]	2014	Contact tracing	Generic	Bluetooth, phone logs	Consent, temporal aggregation
Ferretti et al [43]	2020	Contact tracing	COVID-19	Code scanning, GPS	Consent, transparency
Jeong et al [44]	2019	Contact tracing	Generic	Magnetometer	Geospatial aggregation
Kim et al [45]	2019	Outbreak detection	ILI	Not specified	None
Kim et al [46]	2016	Location-based risk assessment, mobility tracking	Zika	GPS	None
Leal Neto et al [47]	2017	Outbreak detection	Generic	GPS, manual entry	Data minimization
Leal Neto et al [48]	2020	Outbreak detection	Generic	Not specified	Consent
Lwin et al [49]	2014	Outbreak detection	Dengue	GPS, manual entry	None
Michael and Abbas [50]	2020	Contact tracing	COVID-19	Bluetooth	Pseudonymization
Miller et al [51]	2018	Location-based risk assessment, outbreak detection	ILI	GPS	None
Navin et al [52]	2017	Location-based risk assessment, outbreak detection	Generic	Not specified	None
Nguyen et al [53]	2017	Contact tracing	Generic	Magnetometer	Geospatial aggregation
Olson et al [54]	2017	Outbreak detection	Gastroenteritis	Manual entry	None
Okumura [55]	2019	Contact tracing	Generic	GPS, GSM[c]	Data minimization
Prieto et al [56]	2015	Outbreak detection	ILI	Manual entry	None
Rodriguez-Valero et al [57]	2018	Outbreak detection	Zika	GPS	Pseudonymization
Sugiura et al [58]	2010	Outbreak detection	Generic	Manual entry	None
Tripathy et al [59]	2020	Contact tracing	COVID-19	Bluetooth	None
Vazquez-Prokopec et al [60]	2013	Mobility tracking	Generic	GPS	None
Wang et al [61]	2020	Contact tracing	COVID-19	GPS	None
Yasaka et al [20]	2020	Contact tracing	COVID-19	Code scanning	Pseudonymization
Zhang et al [62]	2013	Contact tracing	Generic	Bluetooth	None

[a]ILI: influenza-like illness.

[b]IP: Internet Protocol.

[c]GSM: Global System for Mobile Communications.

As can be seen, more than half of the articles have been published in the last 3 years, indicating a current interest in the topic. In the following sections, we will present further analyses to investigate the questions outlined in the "Objectives" section.

Use Case Dimension

The first objective of this study was to take a broader look at functionalities of mobile health apps using spatial data for infectious disease management.

Within the selected set of articles, we identified four use cases, which we assigned to two distinct categories. The first category is *user-centered services* (ie, solutions focusing on individual health). The first such use case is *automated contact tracing*, where data on contacts between individuals and on infections is used to determine exposure risk and to notify individuals if necessary. The second such use case is *location-based risk assessment*, where the locations of individuals are used to warn them when entering areas with high disease activity. The latter

type of solutions might also be used to estimate individual exposure risk.

The second category is *disease surveillance* (ie, solutions focusing on population health), which involves collecting and analyzing data to monitor the occurrence of diseases within the population with the aim to support public health interventions. The first use cases in this category were *outbreak detection*, which can include syndromic surveillance (eg, based on symptoms reported or data about confirmed cases to determine areas with disease activity). The second such use case is *mobility tracking*, where patterns about the movement of larger groups of individuals are determined to support various analyses, such as the modeling of disease dynamics.

We emphasize that, although these use cases can be implemented independently of one another, there are obvious relationships. For example, data collected via an outbreak detection app could also be used to provide location-based risk assessment services. Moreover, both individual-level as well as population-level use cases ultimately aim to protect the entire population and its individuals. Table 3 illustrates how often the use cases have been described.

Table 3. Overview of use cases covered.

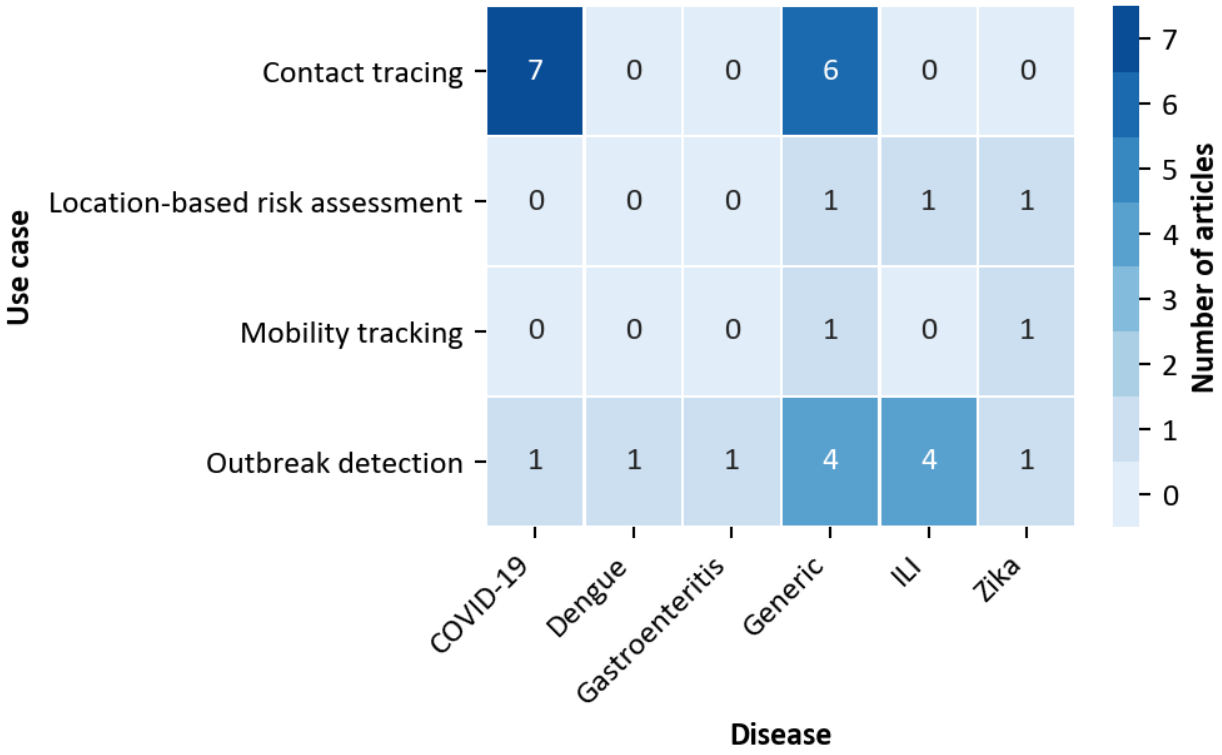
Use case	Articles, n
Contact tracing	13
Outbreak detection	12
Location-based risk assessment	3
Mobility tracking	2

Within the 27 papers, automated contact tracing has been addressed 13 (48.1%) times. Outbreak detection was covered by 12 (44.4%) of the papers. Location-based risk assessment was described in 3 (11.1%) papers, and mobility tracking in 2 (7.4%) papers.

As mentioned before, different use cases can also be combined with each other. However, the articles identified are quite specialized: 24 (89%) focused on a single use case. The remaining 3 (11%) all combined location-based risk assessment with an additional use case.

Current discussions on participatory mobile health solutions in the context of COVID-19 largely focus on automated contact tracing only. To gain insights into potential directions for future developments, we also analyzed which use cases have been described for which disease (see “Disease Dimension” section for more details). Figure 2 shows the common distribution of use cases described and diseases focused on.

Figure 2. Relationships between diseases and use cases described. ILI: Influenza-like illness.



As can be seen, in accordance with current discussions, papers that address COVID-19 focused nearly exclusively on automated contact tracing. For other diseases, however, several further use cases have been described. Moreover, a wide range of generic concepts has been proposed independently of a concrete disease. These focused mainly on automated contact tracing and outbreak detection. We also analyzed in which year solutions for certain use cases have been described in the articles analyzed. An interesting finding is that 69% (9/13) of the papers focusing on automated contact tracing were published in the years 2019 and 2020, while 71% (10/14) of all papers focusing on other use cases were published in 2018 or earlier.

Technology Dimension

In this dimension, we aimed to obtain an overview of technologies that have been proposed to implement the solutions identified. Moreover, we wanted to gather insights into the degree to which technical implementation options and supported functionalities are related to one another.

In total, we identified eight distinct technologies that have been proposed for collecting spatial data: (1) GPS receivers collect absolute coordinates using a satellite system, (2) Bluetooth can be used to estimate the physical proximity between two devices and hence collect relative spatial data (ie, contacts between people), (3) manual entry refers to the manual recording of spatial information such as a living address, (4) magnetometers can be used to measure the small magnetic interferences produced in proximity to a second magnetometer, (5) code scanning refers to the scanning of a code such as a QR code to record a contact or a certain location, (6) Global System for Mobile Communications (GSM) can be used to estimate the absolute geospatial position, (7) Internet Protocol (IP)–based geolocation refers to the estimation of the absolute location by checking the range a user's IP address is assigned to, and (8) phone logs refer to the relative position between people by analyzing phone calling lists and text messages. The frequency with which individual technologies were used is displayed in Table 4.

Table 4. Overview of technologies used for collecting spatial data (three solutions with unspecified technology excluded).

Sensor technology	Articles, n
GPS	10
Bluetooth	7
Manual entry	6
Code scanning	2
Magnetometer	2
Phone logs	1
IP[a] address geolocation	1
GSM[b]	1

[a]IP: Internet Protocol.

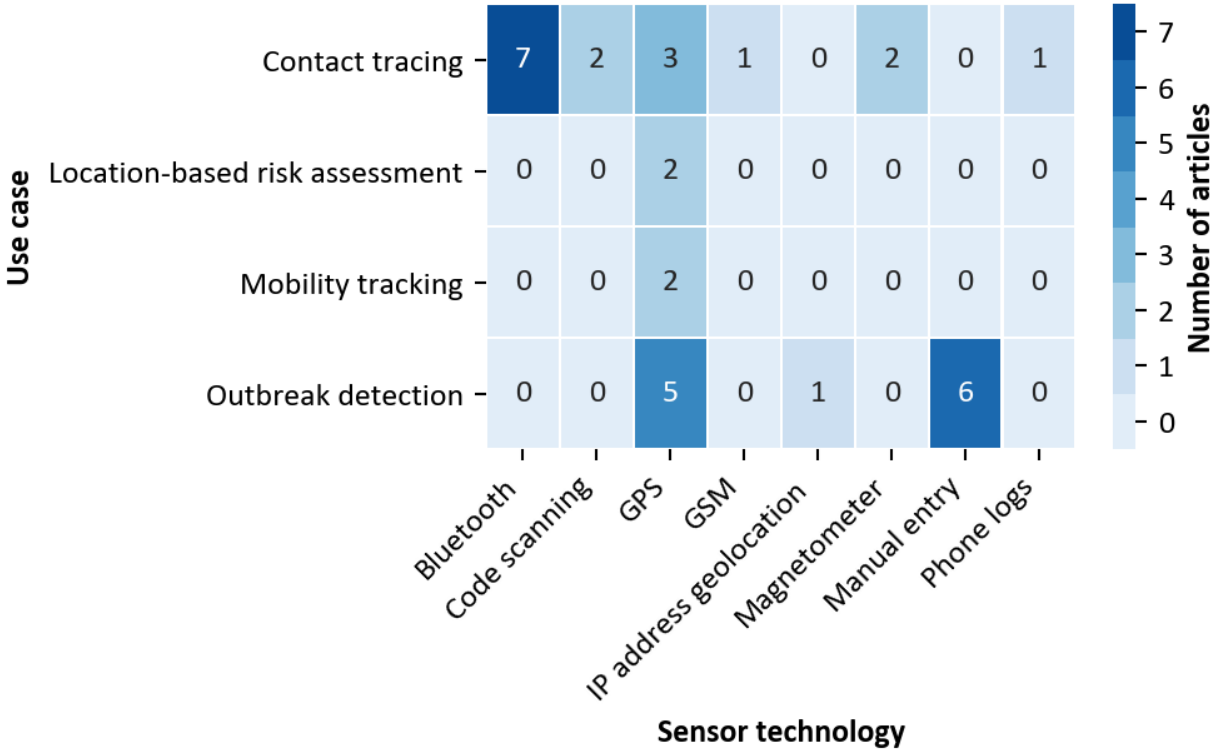
[b]GSM: Global System for Mobile Communications.

Out of 27 articles, GPS, as the most frequent technology, has been described 10 (30.3%) times, followed by Bluetooth, which was mentioned 7 (21.1%) times, and manual entry, which was described in 6 (18.2%) papers. Magnetometers and code scanning were each mentioned in 2 (6.1%) papers. GSM, IP-based geolocation, and phone logs were each suggested in 1 (3.0%) paper. It is noteworthy that 3 articles did not explicitly specify the technology suggested for collecting spatial data (see [45,48,52]). Moreover, it can be noted, that none of the papers described solutions using wireless local area network access points. We also analyzed the ability of the solutions described to track an absolute position (longitude and latitude) or a relative position (proximity to other devices): 60% (n=18) are able to measure absolute positions, while 40% (n=12) can only determine relative positions.

When looking at the time at which papers have been published that use the two most common technologies, GPS and Bluetooth, spikes in the frequency of mentions of Bluetooth can be seen in the years 2013-2014 (the time when the technology first became available on a large scale on commodity devices) and 2020 (the time of the SARS-CoV-2 pandemic). Articles about solutions using GPS have been published continuously over time.

To study how technology might influence the use cases implemented, we analyzed the relationships between both aspects. The result is shown in Figure 3.

Figure 3. Relationship between use cases and technology used to capture spatial data (three solutions with unspecified technology excluded). GSM: Global System for Mobile Communications; IP: Internet Protocol.



It can be seen that Bluetooth is a common technology that has only been used to implement contact tracing (7 papers). Manual data entry is also common but has only been used for outbreak detection (6 papers). GPS, however, is a frequent and versatile technology that has been proposed or implemented for all four use cases identified (12 papers in total). The two remaining technologies that have been suggested more than once are code scanning and magnetometers. Both have only been used to support contact tracing solutions.

Disease Dimension

An important objective of our analysis was to study potential relationships between disease properties and use cases as well as technologies used. We identified five specific diseases addressed: (1) COVID-19 and (2) influenza-like illness (ILI) are mainly transmitted by droplet infection and cause symptoms like coughing and fever; (3) Dengue and (4) Zika are transmitted mainly by mosquitos and cause symptoms like rash, vomiting, and fever; and (5) gastroenteritis is mainly transmitted by smear infection and causes symptoms like diarrhea, vomiting, and fever. As a first step, we analyzed the frequency with which certain diseases have been addressed by publications. The results are shown in Table 5.

Table 5. Overview of diseases targeted by the solutions analyzed.

Disease	Articles, n
Generic	11
COVID-19	8
ILI[a]	4
Zika	2
Dengue	1
Gastroenteritis	1

[a]ILI: influenza-like illness.

Out of 27 articles, 11 (40.7%) of the papers described generic solutions not designed for a specific disease. Of the remaining papers, 8 (29.6%) mentioned COVID-19, 4 (14.8%) mentioned ILI, 2 (7.4%) mentioned Zika, 1 (3.7%) mentioned Dengue, and 1 (3.7%) mentioned gastroenteritis.

We then analyzed whether there was a specific relationship between disease properties and the use cases suggested. For this purpose, we again refer to Figure 3. It can be seen that, obviously, automated contact tracing has only been suggested for highly infectious human-to-human transmissible diseases.

For other diseases that are transmitted by vectors, outbreak detection, mobility tracking, and location-based risk assessment are more common. To study whether a similar relationship can also be found regarding the sensor technology, we compiled the data presented in Figure 4.

Figure 4. Relationship between diseases and sensor technology (three solutions with unspecified technology excluded). ILI: Influenza-like illness; GSM: Global System for Mobile Communications; IP: Internet Protocol.

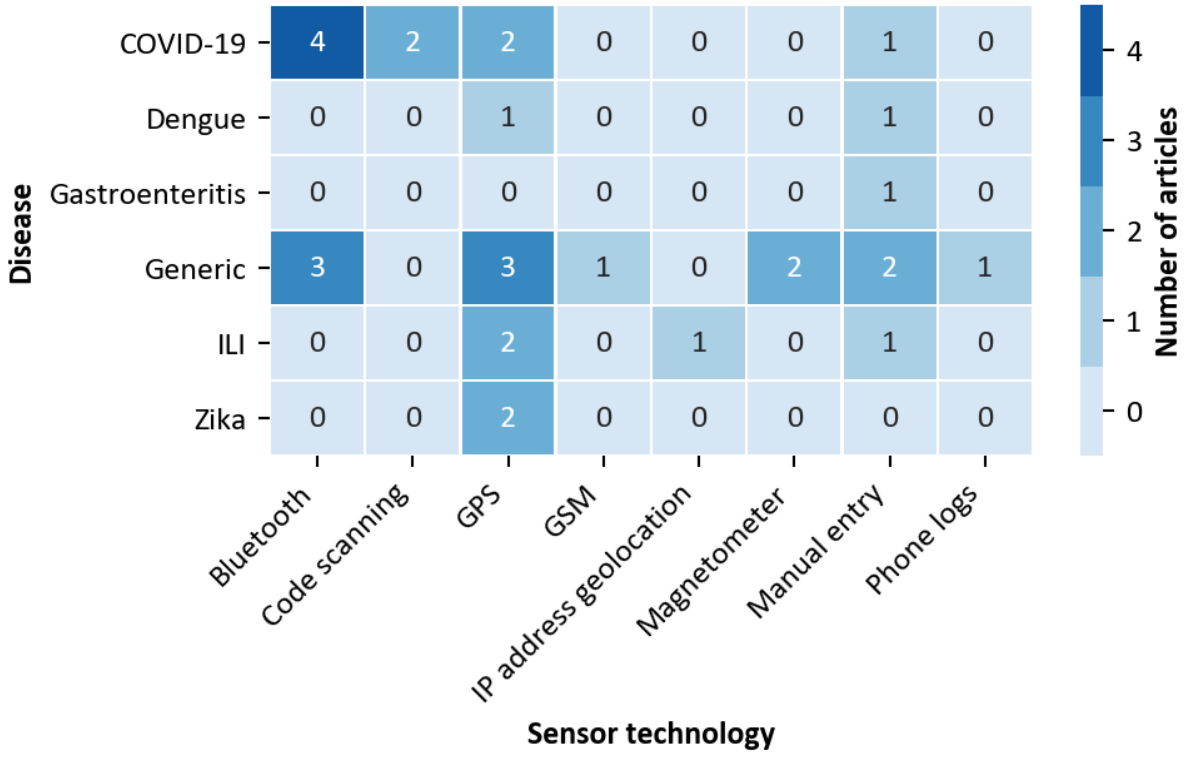


Figure 4 shows that Bluetooth, which is a solution for exclusively tracking relative spatial data (ie, contacts), has only been suggested for generic disease tracking or for managing COVID-19. In contrast, technologies for tracking absolute positions (primarily GPS) have more frequently been proposed for other diseases. It is also worth mentioning that 2 papers suggested the use of code scanning for COVID-19, which is a versatile technology that can be used to track both absolute and relative spatial data.

Data Protection Dimension

Data protection is a central aspect in many discussions on automated contact tracing and related use cases. To gain insights into measures that have been proposed, we first analyzed the protection measures mentioned in the selected articles.

We identified six different data protection measures: (1) pseudonymization, which refers to the replacement of identifying information with random meaningless identifiers; (2) geospatial and (3) temporal aggregation, which refers to techniques for reducing the uniqueness of data; (4) data minimization, which is a privacy-by-design measure implying that as little information as necessary is stored and processed; (5) consent, which means that users are explicitly asked to permit data processing, typically within a study setting; and (6) transparency, which refers to the general principle of communicating which data is stored and how it is processed.

Table 2 shows that 52% (n=14) of the papers did list privacy protection measures, while 48% (n=13) of papers did not mention data protection aspects. When taking a look at the time in which the individual papers were published, it can be seen that there is a trend toward more consideration of data protection aspects in recent years. In total, data protection measures were only mentioned in one of the papers published before 2017.

We then analyzed how often the individual measures were suggested or implemented. The results are shown in Table 6.

Table 6. Data protection measures mentioned in the articles selected.

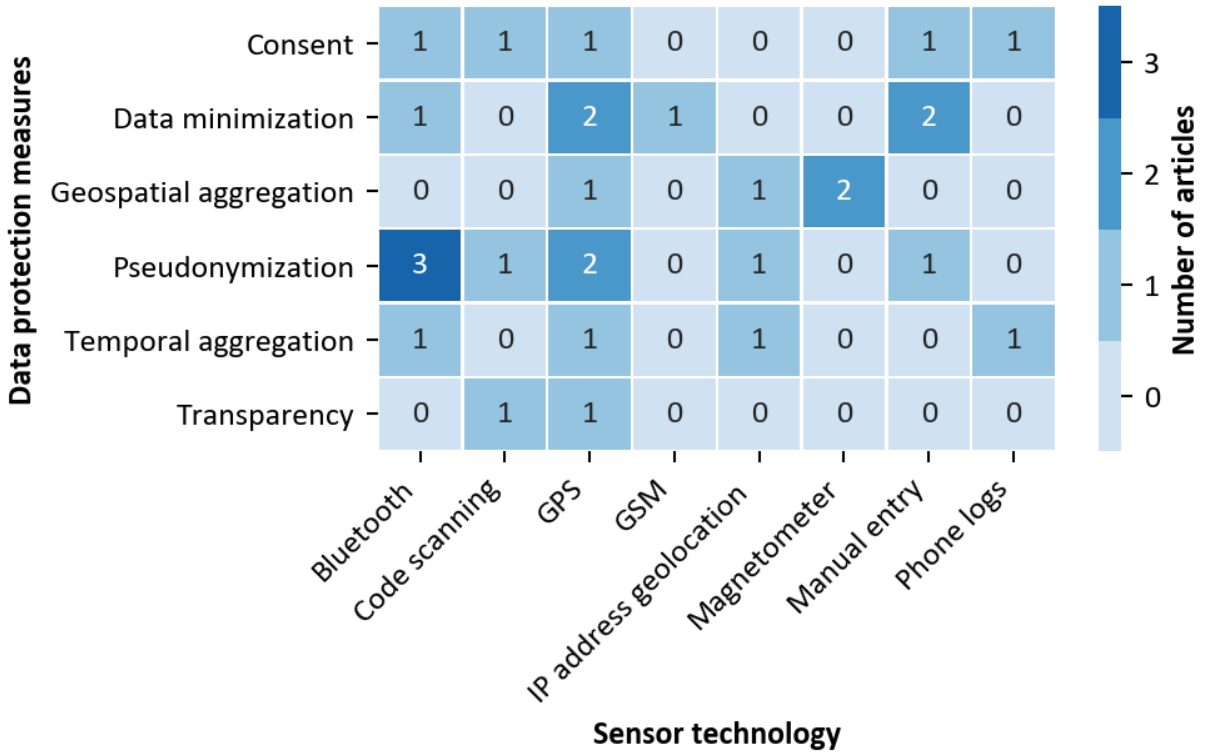
Data protection measure	Articles, n
Pseudonymization	7
Consent	4
Data minimization	4
Geospatial aggregation	3
Temporal aggregation	2
Transparency	1

Table 6 shows that pseudonymization, out of 27 articles, has been addressed 7 (33.3%) times, followed by consent and data minimization, which were each mentioned 4 (19.0%) times. Moreover, geospatial aggregation was described 3 (14.2%) times, temporal aggregation 2 (9.5%) times, and transparency 1 (4.7%) time. These measures can be categorized into organizational and technical measures. The organizational measures transparency and consent have been mentioned 5 (23.8%) times, and the remaining technical measures have been mentioned 16 (76.2%) times.

Regarding the relationships between use cases and data protection measures, it must be noted that all 4 use cases identified were potentially privacy-invading. However, data protection measures have not been mentioned in any of the papers covering location-based risk assessment or mobility tracking.

Finally, we also studied the relationship between technologies and protection measures proposed. The results are shown in Figure 5.

Figure 5. Sensor technologies utilized and protection methods mentioned (three solutions with unspecified technology excluded) GSM: Global System for Mobile Communications; IP: Internet Protocol.



It can be seen that many different protection measures have been mentioned in papers focusing on the popular GPS sensor technology. Several measures have also been suggested for protecting Bluetooth-based solutions including pseudonymization, which is particularly important in this context and implemented by exchanging random, pseudonymous identifiers. Most other mentions of protection measures are used in papers focusing on various types of sensors for absolute spatial data other than GPS.

Discussion

Use Case Dimension

To broaden the discussion on technical solutions to combat COVID-19 and to propose extension to their functionalities, we first focused on identifying specific use cases. Although most recent publications focus on automated contact tracing, we have identified three additional use cases. Use cases can be grouped

into two categories: user-centered services and disease surveillance. All these functionalities have already been discussed and used in the context of COVID-19 [15,16] but in an isolated manner. Future work could, therefore, focus on developing concepts and solutions that support different use cases in a combined manner. Our data also indicates, however, that this might be challenging: 89% (24/27) of all papers analyzed focus exclusively on one use case. Among the notable exceptions is a solution combining outbreak detection with contact tracing for COVID-19 [63].

There might be additional aspects to consider. First, different use cases might be particularly important for different infectious diseases, and the pandemic scale of COVID-19 might be an important factor. We investigated this further in the disease dimension. Second, there might also be technical challenges associated with supporting different use cases and protecting privacy in this process. We investigated this in the technology and data protection dimensions.

Technology Dimension

We found eight different technologies in the selected literature from which GPS and Bluetooth were most frequently suggested. This is consistent with current discussions on contact tracing apps, which seem to be dominated by BLE. Most technologies were only proposed within the context of a specific use case due to their differences in supporting the measurement of spatial data. For example, Bluetooth has only been proposed for contact tracing, as this technology is only suited for detecting the proximity between individual devices. Moreover, manual data entry has only been suggested for outbreak detection, as it is only suited for tracking static locations such as living addresses. Our results also show that GPS is a highly versatile technology, which has been proposed as a basis for all use cases considered in at least 1 paper. This makes sense, as GPS can be used to collect dynamically absolute locations, which can also be used to determine (to some degree) relative and static spatial data. However, a major challenge arising from this is privacy, which we cover in the data protection dimension.

Our results also suggest that it might be worthwhile to combine different technologies to implement multiple use cases. For example, Bluetooth could be used to implement automated contact tracing, combined with GPS receivers to support location-based risk assessment and mobility tracking. Moreover, this could be combined with manual entry of symptoms or integrated with wearables to collect further data on disease spread. An important prerequisite might be, however, that adequate privacy-enhancing technologies are implemented.

Disease Dimension

Our results showed that COVID-19 is already the most frequently covered disease in papers on automated contact tracing or related use cases for managing infectious diseases. Obviously, this is due to the scale and global consequences of the COVID-19 pandemic. Moreover, our data shows that most solutions for managing COVID-19 have suggested Bluetooth as the main technology for collecting spatial data. Apart from privacy considerations, this is also well justified by disease properties: COVID-19 is highly infectious, droplet infection is an important route, and there can be a long period of infectiousness before symptom onset. Thus, a technology is needed that can detect close proximity between people at scale with relatively high accuracy [64]. With diseases like Zika and Dengue that are vector-borne, technologies are needed that can capture absolute spatial data (eg, to identify clusters of infections). Code scanning is a technology that can measure absolute and relative spatial data and may, therefore, also be suited for managing diseases like COVID-19. However, code scanning is cumbersome to use, as it is nonautomated and may, therefore, not be well received by the public. It was noticeable that we did not find any articles on several common infectious diseases such as malaria or HIV. The reason is, as we demonstrated, that disease properties do have an influence on the adequacy of solutions to manage them, and solutions focused on these diseases fall out of the scope of this study. One example are apps to estimate the size of local mosquito populations, which are important for managing malaria but typically use special equipment such as mosquito traps [65].

Data Protection Dimension

Our results show that various types of organizational and technical safeguards have been suggested for protecting privacy when implementing citizen-centered disease surveillance apps. Moreover, more emphasis has been put on this aspect in recent years. Due to the scale of the current pandemic, the topic has become even more important.

There are two use cases that we identified (ie, location-based risk assessment and mobility tracking) for which no protection measures have been proposed in the literature investigated. This is surprising, as both can be considered potentially privacy-invading. Future work will be needed to integrate such use cases into current participatory surveillance apps in a privacy-conscious manner.

Regarding relationships between privacy protection measures and technologies used, we found that more measures have been suggested for technologies that are widely employed and whose sensors provide absolute positions. Our data shows that pseudonymization is an important measure for Bluetooth-based solutions and, thus, especially for the contact tracing use case.

Finally, we would like to add that several techniques have only been mentioned in papers describing concepts and might have been suggested with a rather naïve view on the topic. Most importantly, it is well-known that geospatial as well as temporal aggregation are challenging to implement in a manner offering a high degree of protection [66]. Moreover, even if such an implementation can be developed, it might have significant impacts on the precision of location data [67]. At the same time, solutions focusing on relative spatial data combined with pseudonymization and implemented without continuous data exchange with central services such as the German Corona-Warn-App [31] are considered the current “gold standard” in privacy-preserving automated contact tracing. Future solutions combining relative with absolute spatial data to support further use cases might be built using alternative approaches such as secure multiparty computing protocols [68] and differential privacy [69]. Finally, it is worth mentioning that data protection is not only a technical but also a social

challenge, as it is highly connected with public reception and trust.

Conclusion

In this paper, we have analyzed the literature and discovered several relationships between disease properties, use cases, and technologies. To our knowledge, many of these general dependencies have not been described previously. We, therefore, believe that our results can help with enhancing current solutions for contact tracing and related use cases, and with developing novel, more comprehensive concepts. In addition, the described dependencies could support bottom-up development processes leading to solutions that are more likely to stand the test in real-world implementations. Moreover, we have studied data protection measures that have been suggested and discussed their suitability for different technical environments and use cases. We believe that it will be necessary to employ innovative privacy-enhancing technologies to build comprehensive solutions offering additional functionalities such as population surveillance or individual alerting while maintaining the privacy of citizens.

In the future, when more data on the implementation of solutions in the context of the COVID-19 pandemic will become available, we plan to investigate relationships between their properties along the axes considered in this study and outcomes achieved within specific patient populations.

Authors' Contributions

FP und FW initiated and conceptualized the work. All authors collected and analyzed the data. FP and FW drafted the manuscript. MJ and TM revised the manuscript. All authors read and approved the final manuscript

Conflicts of Interest

None declared.

Multimedia Appendix 1
Selection process and collected data.
[XLSX File (Microsoft Excel File), 194 KB - mhealth_v8i11e22594_app1.xlsx]

References

1. Report of the WHO-China Joint Mission on coronavirus disease 2019 (COVID-19). World Health Organization. 2020. URL: https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf
2. Lavezzo E, Franchin E, Ciavarella C, Cuomo-Dannenburg G, Barzon L, Del Vecchio C, Imperial College London COVID-19 Response Team, et al. Suppression of COVID-19 outbreak in the municipality of Vo, Italy. medRxiv Preprint posted online April 18, 2020. [doi: 10.1101/2020.04.17.20053157]
3. Flaxman S, Mishra S, Gandy A, Unwin HJT, Coupland H, Mellan TA, Imperial College London COVID-19 Response Team, et al. Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in European countries: technical description update. arXiv Preprint posted online April 23, 2020.
4. Chu D, Akl E, Duda S, Solo K, Yaacoub S, Schünemann HJ, COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 2020 Jun 27;395(10242):1973-1987 [FREE Full text] [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
5. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science 2020 Apr 24;368(6489):395-400 [FREE Full text] [doi: 10.1126/science.aba9757] [Medline: 32144116]
6. Fernandes N. Economic effects of coronavirus outbreak (COVID-19) on the world economy. SSRN J Preprint posted online March 23, 2020. [doi: 10.2139/ssrn.3557504]
7. Nussbaumer-Streit B, Mayr V, Dobrescu A, Chapman A, Persad E, Klerings I, et al. Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. Cochrane Database Syst Rev 2020 Apr 08;4:CD013574. [doi: 10.1002/14651858.CD013574] [Medline: 32267544]
8. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): a review. Int J Surg 2020 Jun;78:185-193 [FREE Full text] [doi: 10.1016/j.ijsu.2020.04.018] [Medline: 32305533]
9. Jensen JB. Economic consequences of the COVID-19 pandemic. Copenhagen Economics. 2020 Mar. URL: https://www.copenhageneconomics.com/dyn/resources/Publication/publicationPDF/0/530/1585835646/copenhagen-economics_economic-consequences-covid-19.pdf
10. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry 2020 Jun;7(6):547-560. [doi: 10.1016/s2215-0366(20)30168-1]

11. Coronavirus business impact: evolving perspective. McKinsey & company. 2020. URL: https://www.mckinsey.com/business-functions/risk/our-insights/covid-19-implications-for-business [accessed 2020-07-17]
12. Lee V, Chiew C, Khong W. Interrupting transmission of COVID-19: lessons from containment efforts in Singapore. J Travel Med 2020 May 18;27(3) [FREE Full text] [doi: 10.1093/jtm/taaa039] [Medline: 32167146]
13. Ferguson NM, Cummings DAT, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. Nature 2006 Jul 27;442(7101):448-452 [FREE Full text] [doi: 10.1038/nature04795] [Medline: 16642006]
14. Osterholm MT, Hedberg CW. Epidemiologic principles. In: Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases. Amsterdam: Elsevier; 2015:146-158.
15. 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]
16. Kamel Boulos MN, Geraghty EM. Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. Int J Health Geogr 2020 Mar 11;19(1):8 [FREE Full text] [doi: 10.1186/s12942-020-00202-8] [Medline: 32160889]
17. De Carli A, Franco M, Gassmann A, Killer C, Rodrigues B, Scheid E, et al. WeTrace -- a privacy-preserving mobile COVID-19 tracing approach and application. arXiv Preprint posted online April 19, 2020. [FREE Full text]
18. Cho H, Ippolito D, Yu YW. Contact tracing mobile apps for COVID-19: privacy considerations and related trade-offs. arXiv Preprint posted online March 25, 2020. [FREE Full text]
19. Xia Y, Lee G. How to return to normalcy: fast and comprehensive contact tracing of COVID-19 through proximity sensing using mobile devices. arXiv Preprint posted online April 5, 2020. [FREE Full text]
20. 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]
21. 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. 2020 May 07. URL: https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker/ [accessed 2020-07-17]
22. 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 May 8, 2020. [doi: 10.1101/2020.05.05.20091587]
23. 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]
24. Abeler J, Bäcker M, Buermeyer U, Zillessen H. COVID-19 contact tracing and data protection can go together. JMIR mHealth uHealth 2020 Apr 20;8(4):e19359 [FREE Full text] [doi: 10.2196/19359] [Medline: 32294052]
25. 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]
26. Chen H, Yang B, Pei H, Liu J. Next generation technology for epidemic prevention and control: data-driven contact tracking. IEEE Access 2019;7:2633-2642. [doi: 10.1109/access.2018.2882915]
27. Segal E, Zhang F, Lin X, King G, Shalem O, Shilo S, et al. Building an international consortium for tracking coronavirus health status. Nat Med 2020 Aug;26(8):1161-1165. [doi: 10.1038/s41591-020-0929-x] [Medline: 32488218]
28. Pan-European Privacy-Preserving Proximity Tracing. 2020. URL: http://www.pepp-pt.org/
29. DP-3T / documents. GitHub. 2020. URL: https://github.com/DP-3T/documents
30. Whitepaper: slowing the spread of infectious diseases using crowdsourced data. Covid Watch. 2020. URL: https://www.covid-watch.org/covid_watch_whitepaper.pdf
31. Open-Source Project Corona-Warn-App. 2020. URL: https://www.coronawarn.app/en/ [accessed 2020-07-17]
32. Ding X, Clifton D, JI N, Lovell NH, Bonato P, Chen W, et al. Wearable sensing and telehealth technology with potential applications in the coronavirus pandemic. IEEE Rev Biomed Eng 2020 May 11. [doi: 10.1109/rbme.2020.2992838]
33. Ming DK, Sangkaew S, Chanh HQ, Nhat PT, Yacoub S, Georgiou P, et al. Continuous physiological monitoring using wearable technology to inform individual management of infectious diseases, public health and outbreak responses. Int J Infect Dis 2020 Jul;96:648-654 [FREE Full text] [doi: 10.1016/j.ijid.2020.05.086] [Medline: 32497806]
34. Tom-Aba D, Nguku PM, Arinze CC, Krause G. Assessing the concepts and designs of 58 mobile apps for the management of the 2014-2015 West Africa Ebola outbreak: systematic review. JMIR Public Health Surveill 2018 Oct 29;4(4):e68 [FREE Full text] [doi: 10.2196/publichealth.9015] [Medline: 30373727]
35. Ouedraogo B, Gaudart J, Dufour J. How does the cellular phone help in epidemiological surveillance? A review of the scientific literature. Inform Health Soc Care 2019 Jan;44(1):12-30. [doi: 10.1080/17538157.2017.1354000] [Medline: 28829691]
36. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143 [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]

37. Tricco A, 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]
38. Abbas R, Michael K. COVID-19 contact trace app deployments: learnings From Australia and Singapore. IEEE Consumer Electron Mag 2020 Sep 1;9(5):65-70. [doi: 10.1109/mce.2020.3002490]
39. Ackley S, Pilewski S, Petrovic V, Worden L, Murray E, Porco T. Assessing the utility of a smart thermometer and mobile application as a surveillance tool for influenza and influenza-like illness. Health Informatics J 2020 Sep;26(3):2148-2158 [FREE Full text] [doi: 10.1177/1460458219897152] [Medline: 31969046]
40. Barrat A, Cattuto C, Tozzi A, Vanhems P, Voirin N. Measuring contact patterns with wearable sensors: methods, data characteristics and applications to data-driven simulations of infectious diseases. Clin Microbiol Infect 2014 Jan;20(1):10-16. [doi: 10.1111/1469-0691.12472] [Medline: 24267942]
41. Chan A, Drew D, Nguyen L, Joshi A, Ma W, Guo C, COPE Consortium. The COronavirus Pandemic Epidemiology (COPE) Consortium: a call to action. Cancer Epidemiol Biomarkers Prev 2020 Jul;29(7):1283-1289 [FREE Full text] [doi: 10.1158/1055-9965.EPI-20-0606] [Medline: 32371551]
42. Farrahi K, Emonet R, Cebrian M. Epidemic contact tracing via communication traces. PLoS One 2014;9(5):e95133. [doi: 10.1371/journal.pone.0095133] [Medline: 24787614]
43. 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) [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
44. Jeong S, Kuk S, Kim H. A smartphone magnetometer-based diagnostic test for automatic contact tracing in infectious disease epidemics. IEEE Access 2019;7:20734-20747. [doi: 10.1109/access.2019.2895075]
45. Kim M, Yune S, Chang S, Jung Y, Sa S, Han H. The fever coach mobile app for participatory influenza surveillance in children: usability study. JMIR mHealth uHealth 2019 Oct 17;7(10):e14276 [FREE Full text] [doi: 10.2196/14276] [Medline: 31625946]
46. Kim M, Lee JY, Kim H. Warning and detection system for epidemic disease. 2016 Presented at: 2016 International Conference on Information and Communication Technology Convergence (ICTC); 2016; Jeju, South Korea p. 478-483. [doi: 10.1109/ictc.2016.7763517]
47. Leal Neto O, Dimech GS, Libel M, de Souza WV, Cesse E, Smolinski M, et al. Saúde na Copa: the world's first application of participatory surveillance for a mass gathering at FIFA World Cup 2014, Brazil. JMIR Public Health Surveill 2017 May 04;3(2):e26 [FREE Full text] [doi: 10.2196/publichealth.7313] [Medline: 28473308]
48. Leal Neto O, Cruz O, Albuquerque J, Nacarato de Sousa M, Smolinski M, Pessoa Cesse E, et al. Participatory surveillance based on crowdsourcing during the Rio 2016 Olympic Games using the Guardians of Health platform: descriptive study. JMIR Public Health Surveill 2020 Apr 07;6(2):e16119 [FREE Full text] [doi: 10.2196/16119] [Medline: 32254042]
49. Lwin MO, Vijaykumar S, Fernando ONN, Cheong SA, Rathnayake VS, Lim G, et al. A 21st century approach to tackling dengue: crowdsourced surveillance, predictive mapping and tailored communication. Acta Trop 2014 Feb;130:100-107. [doi: 10.1016/j.actatropica.2013.09.021] [Medline: 24161879]
50. Michael K, Abbas R. Behind COVID-19 contact trace apps: the Google–Apple partnership. IEEE Consumer Electron Mag 2020 Sep 1;9(5):71-76. [doi: 10.1109/mce.2020.3002492]
51. Miller A, Singh I, Koehler E, Polgreen P. A smartphone-driven thermometer application for real-time population- and individual-level influenza surveillance. Clin Infect Dis 2018 Jul 18;67(3):388-397. [doi: 10.1093/cid/ciy073] [Medline: 29432526]
52. Navin K, Mukesh Krishnan MB, Lavanya S, Shanthini A. A mobile health based smart hybrid epidemic surveillance system to support epidemic control programme in public health informatics. 2017 Presented at: 2017 International Conference on IoT and Application (ICIOT); 2017; Nagapattinam, India. [doi: 10.1109/iciota.2017.8073606]
53. Nguyen KA, Watkins C, Luo Z. Co-location epidemic tracking on London public transports using low power mobile magnetometer. 2017 Presented at: 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN); 2017; Sapporo, Japan. [doi: 10.1109/ipin.2017.8115963]
54. Olson D, Lamb M, Lopez MR, Colborn K, Paniagua-Avila A, Zacarias A, et al. Performance of a mobile phone app-based participatory syndromic surveillance system for acute febrile illness and acute gastroenteritis in rural Guatemala. J Med Internet Res 2017 Nov 09;19(11):e368 [FREE Full text] [doi: 10.2196/jmir.8041] [Medline: 29122738]
55. Okumura T. Tracing infectious agents with mobile location information: a simple and effective countermeasure against epidemic risks. 2019 Presented at: 2019 IEEE Global Humanitarian Technology Conference (GHTC); 2019; Seattle, WA. [doi: 10.1109/ghtc46095.2019.9033023]
56. Prieto JT, Jara JH, Alvis JP, Furlan LR, Murray CT, Garcia J, et al. Will participatory syndromic surveillance work in Latin America? Piloting a mobile approach to crowdsource influenza-like illness data in Guatemala. JMIR Public Health Surveill 2017 Nov 14;3(4):e87 [FREE Full text] [doi: 10.2196/publichealth.8610] [Medline: 29138128]
57. Rodriguez-Valero N, Luengo Oroz M, Cuadrado Sanchez D, Vladimirov A, Espriu M, Vera I, et al. Mobile based surveillance platform for detecting Zika virus among Spanish Delegates attending the Rio de Janeiro Olympic Games. PLoS One 2018;13(8):e0201943 [FREE Full text] [doi: 10.1371/journal.pone.0201943] [Medline: 30133492]

58. Sugiura H, Ohkusa Y, Akahane M, Sugahara T, Okabe N, Imamura T. Construction of syndromic surveillance using a web-based daily questionnaire for health and its application at the G8 Hokkaido Toyako Summit meeting. Epidemiol Infect 2010 Oct;138(10):1493-1502. [doi: 10.1017/S095026880999149X] [Medline: 20067657]
59. Tripathy A, Mohapatra A, Mohanty S, Kougianos E, Joshi A, Das G. EasyBand: a wearable for safety-aware mobility during pandemic outbreak. IEEE Consumer Electron Mag 2020 Sep 1;9(5):57-61. [doi: 10.1109/mce.2020.2992034]
60. Vazquez-Prokopec G, Bisanzio D, Stoddard S, Paz-Soldan V, Morrison A, Elder J, et al. Using GPS technology to quantify human mobility, dynamic contacts and infectious disease dynamics in a resource-poor urban environment. PLoS One 2013;8(4):e58802 [FREE Full text] [doi: 10.1371/journal.pone.0058802] [Medline: 23577059]
61. Wang S, Ding S, Xiong L. A new system for surveillance and digital contact tracing for COVID-19: spatiotemporal reporting over network and GPS. JMIR mHealth uHealth 2020 Jun 10;8(6):e19457 [FREE Full text] [doi: 10.2196/19457] [Medline: 32499212]
62. Zhang Z, Wang H, Lin X, Fang H, Xuan D. Effective epidemic control and source tracing through mobile social sensing over WBANs. 2013 Presented at: 2013 Proceedings IEEE INFOCOM; 2013; Turin, Italy. [doi: 10.1109/infcom.2013.6566783]
63. Lewis DM. Frequently asked questions. CoEpi: Community Epidemiology in Action. 2020. URL: https://www.coepi.org/faq/
64. Leith DJ, Farrell S. Coronavirus contact tracing: evaluating the potential of using bluetooth received signal strength for proximity detection. SIGCOMM Comput Commun Rev 2020 Oct 26;50(4):66-74. [doi: 10.1145/3431832.3431840]
65. Braz Sousa L, Fricker SR, Doherty SS, Webb CE, Baldock KL, Williams CR. Citizen science and smartphone e-entomology enables low-cost upscaling of mosquito surveillance. Sci Total Environ 2020 Feb 20;704:135349. [doi: 10.1016/j.scitotenv.2019.135349] [Medline: 31837870]
66. de Montjoye YA, Hidalgo C, Verleysen M, Blondel V. Unique in the crowd: the privacy bounds of human mobility. Sci Rep 2013;3:1376. [doi: 10.1038/srep01376] [Medline: 23524645]
67. Zang H, Bolot J. Anonymization of location data does not work: a large-scale measurement study. 2011 Presented at: Mobicom'11: The 17th Annual International Conference on Mobile Computing and Networking; 2011; Las Vegas, NV. [doi: 10.1145/2030613.2030630]
68. Reichert L, Brack S, Scheuermann B. Privacy-preserving contact tracing of COVID-19 patients. IACR Cryptol ePrint Arch 2020.
69. Aktay A, Bavadekar S, Cossoul G, Davis J, Desfontaines D, Fabrikant A, et al. Google COVID-19 community mobility reports: anonymization process description. arXiv Preprint posted online April 8, 2020. [FREE Full text]

Abbreviations

BLE: Bluetooth Low Energy
GSM: Global System for Mobile Communications
ILI: influenza-like illness
IP: Internet Protocol
NFC: near-field communication
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for scoping reviews
QR: Quick Response
RFID: radio-frequency identification

Edited by G Eysenbach, L Buis; submitted 17.07.20; peer-reviewed by J Jose, X Wu; comments to author 24.08.20; revised version received 26.08.20; accepted 09.10.20; published 10.11.20.

Please cite as:
Wirth FN, Johns M, Meurers T, Prasser F
Citizen-Centered Mobile Health Apps Collecting Individual-Level Spatial Data for Infectious Disease Management: Scoping Review
JMIR Mhealth Uhealth 2020;8(11):e22594
URL: http://mhealth.jmir.org/2020/11/e22594/
doi:10.2196/22594
PMID:33074833

Viewpoint

COVID-19 Contact Tracing Apps: A Technologic Tower of Babel and the Gap for International Pandemic Control

Li Du[1], MBBS, LLB, PhD; Vera Lúcia Raposo[1,2], BA, LLB, LLM, PhD; Meng Wang[1], LLB, LLM

[1]Faculty of Law, University of Macau, Macau, SAR, China

[2]Faculty of Law, University of Coimbra, Coimbra, Portugal

Corresponding Author:
Li Du, MBBS, LLB, PhD
Faculty of Law
University of Macau
Avenida da Universidade, Taipa
Macau, SAR, 999078
China
Phone: 853 88224733
Email: stephendu@um.edu.mo

Related Articles:

Comment in: https://mhealth.jmir.org/2021/5/e26218

Comment in: https://mhealth.jmir.org/2021/5/e26630

This is a corrected version. See correction statement: https://mhealth.jmir.org/2020/12/e26239

Abstract

As the world struggles with the new COVID-19 pandemic, contact tracing apps of various types have been adopted in many jurisdictions for combating the spread of the SARS-CoV-2 virus. However, even if they are successful in containing the virus within national borders, these apps are becoming ineffective as international travel is gradually resumed. The problem rests in the plurality of apps and their inability to operate in a synchronized manner, as well as the absence of an international entity with the power to coordinate and analyze the information collected by the disparate apps. The risk of creating a useless Tower of Babel of COVID-19 contact tracing apps is very real, endangering global health. This paper analyzes legal barriers for realizing the interoperability of contact tracing apps and emphasizes the need for developing coordinated solutions to promote safe international travel and global pandemic control.

(JMIR Mhealth Uhealth 2020;8(11):e23194) doi:10.2196/23194

KEYWORDS

COVID-19; contact tracing apps; privacy; public health; global health

Background

As the novel SARS-CoV-2 virus spreads worldwide—enabled by an unknown number of asymptomatic carriers—individuals continue to be at risk for potentially fatal infections. An effective method to contain the spread of the virus is tracing the movement of confirmed and suspected COVID-19 cases, as well as their close contacts [1]. To this end, many countries have used contact tracing mobile apps for controlling the COVID-19 epidemic. These apps are based on various techniques that can provide users with infection risk-level information specific to their communities and notify users who have been exposed to COVID-19, thus facilitating the public, social organizations, and the government to better prevent and control the epidemic [2].

With countries gradually reopening their borders for business and tourism, the effectiveness of using such apps for containing the spread of the pandemic has greatly weakened. The region-based development of contact tracing apps, along with the distinguished data protection laws used in different countries and regions, has resulted in the disconnection between contact tracing apps. These isolated contact tracing apps, which work

as a technologic Tower of Babel, have lost their abilities to trace and monitor the spread of the pandemic in cross-border travel, triggering challenges for global pandemic control while international travel is resumed.

This viewpoint aims to raise global awareness on the urgent need for establishing a data sharing and transferring mechanism at the international level. It advocates that a consistent international effort should be devoted to developing an international data sharing platform for global pandemic control and removing the legal barriers that hinder the interoperability of contact tracing app technology. All these efforts are aimed to contribute a better response to potential future waves of COVID-19 and future global pandemic crises.

Region-Based Contact Tracing App Development

COVID-19 contact tracing apps are designed and developed at the regional level, employing various technologies [3]. Asian countries, such as China, South Korea, and Singapore, are among the first group worldwide to develop and use tracking measures for fighting against the COVID-19 epidemic. In China for instance, the national government affairs service began developing the Epidemic Prevention and Health Information Code (EPHIC) after the pandemic surge in early February 2020 [4]. After a smart phone scans the EPHIC QR code, users submit their basic health status, residential addresses, and information about whether they have interacted with any confirmed or suspected individuals with COVID-19 within the past 14 days. With the spread of the COVID-19 pandemic, contact tracing apps have been widely used in many other countries.

Australia and New Zealand are also among the first countries that developed a national contact tracing app. For example, the COVIDSafe app launched by the Australian Department of Health helps state and territory health officials to identify any contact app users who have had close contact with patients with COVID-19 [5].

In Canada, contact tracing apps were first developed by provincial governments and used within their own territories. For example, the ABTraceTogether app in Alberta and the BC COVID-19 Support app in British Columbia both leverage epidemic tracking information provided by Canadian health departments to inform users about their epidemic exposure risks. The federal government also developed the COVID Alert app, which help users to better understand the risk levels of their own regions, as well as harness geographic location information to inform other users as to whether they have encountered anyone infected with SARS-CoV-2 [6].

Some states in the United States have developed contact tracing apps for local residents, such as the Care19 app in South Dakota and the Healthy Together app in Utah and Florida. Technology companies and research institutions have also taken up the challenge and have jointly developed software, such as the COVIDWatch app, NOVID app, and the Private Kit app. Through different technical means, such as self-reporting by users, Bluetooth, and GPS data, these apps can detect and record the contacts between app users, provide an indication of the risk of infection, and archive the data for the purpose of epidemic contact tracing [7]. In May 2020, Google and Apple jointly developed contact tracing technology and released application programming interfaces (APIs) that allow governments to work on developing their own contact tracing apps. Since then, more than one-fifth of the states in the United States have used the Google-Apple API code to develop their own contact tracing apps. To further promote the use of contact tracing technology, in September 2020, Google and Apple announced that they will incorporate their contact tracing tool into the latest operating system of their respective smartphones. This new system allows more public health authorities in the United States to embrace contact tracing functions, as no extra burdens are required to develop their own contact tracing apps [8].

In Europe, more than 20 countries have developed contact tracing apps for improving the management of COVID-19 [9]. Some countries, including France, Norway, and Hungary, have created their own model for contact tracing apps, while other countries, such as Germany, Netherland, and Switzerland, have used the Google-Apple API code to develop their contact tracing apps [10]. In June 2020, the UK government announced that it abandoned a centralized contact tracing app and adopted the new Google-Apple framework for future COVID-19 contact tracing software development to help achieve more effective pandemic control [11]. Initially, contact tracing apps were not welcomed by the European population or legal community, mostly due to privacy concerns and the requirements of the General Data Protection Regulation (GDPR). However, the second pandemic wave in Europe moderated the critics and highlighted the potential of contact tracing to control the progression of the virus.

The Formation of the Technologic Tower of Babel

Different Forms of Contact Tracing

Region-based contact tracing apps are developed by using diverse techniques (Table 1) and are used to collect different types of data. For instance, South Korea, New Zealand, Israel, and Taiwan have apps that transmit the users' location. This technology model can allow the data subject to be identified by third parties, as demonstrated by an experience in South Korea, during which the disclosure of an infected individual's location made it possible to identify the person in question [12]. Some governments force the installation of contact apps to monitor the flow of confirmed cases and travelers, while other jurisdictions, such as those in Singapore, Austria, and France, allow users to voluntarily choose to install the apps; users can also decide whether or not to have their data collected (ie, automatically collected via the app or by means of a voluntary act, such as reporting personal health information or by registering their location) [13]. Importantly, domestic and regional laws have different rules on data protection and privacy issues, which hamper transnational data sharing and exchange.

Table 1. Technologies used in COVID-19 contact tracing apps and potential legal challenges.

Technologies used for contact tracing apps	App examples, app name (country/region of origin)	Potential legal challenges
Location tracking by GPS	• Self-quarantine safety protection (South Korea) • Epidemic Prevention and Health Information Code (China) • National Government Service Platform (China) • HaMagen (Israel) • NZ COVID Tracer (New Zealand) • COVID Safe Paths (United States) • Healthy Together (Utah, United States) • Care19 (North Dakota, United States)	• Personal privacy issues, including the disclosure of personal privacy location information, personal health information, etc [14]. • It is not clear who will bear the responsibility if the users' location is not accurate, the disease condition is reported incorrectly due to location information error, and there is a false report. • If the state or the authorities adopt the technology on a large scale, but the technology fails due to satellite signals or technical failures, how can it be remedied? • Some software shares data with the authorities, which could make it possible for the government to access personal location information. The issue is especially sensitive when it comes to the location of people of different races. • Location can reveal a lot of personal information (eg, sexual orientation, religion, political affiliation) that is not directly related to pandemic control. • With regard to apps that keep collected data in a central remote server, if the server is attacked, there will be a massive privacy breach.
Contact tracing by GPS	• Epidemic Prevention and Health Information Code (China) • National Government Service Platform (China) • HaMagen (Israel) • NZ COVID Tracer (New Zealand) • COVID Safe Paths (United States) • Healthy Together (Utah, United States) • Care19 (North Dakota, United States)	• All legal challenges for using GPS to track users' location. • Using GPS for contact tracing requires analyses of the location information of multiple users by means of big data. This involves methods for properly storing and using the personal information of users and other legal issues related to personal data privacy protection.
Contact tracing by Bluetooth	• TraceTogether (Singapore) • COVIDSafe (Australia) • Stopp Corona (Austria) • ABTraceTogether (Alberta, Canada) • COVID Alert (Canada) • ProteGO (Poland) • Corona-Warn-App (Germany) • SwissCovid (Switzerland) • HaMagen (Israel) • COVID Safe Paths (United States) • Healthy Together (Utah, United States) • Care19 (North Dakota, United States) • CovidWatch (United States) • NOVID (United States)	• Although most of the software using Bluetooth technology has claimed that they will not obtain user information and will only warn people about the risk of disease through distance perception, data security is a big concern, as hackers may attack the Bluetooth firmware to obtain the user's personal information and location data information [14]. • Bluetooth technology faces many technical incompatibilities between devices, which will lead to incomplete information collection and, to some extent, the inability to effectively control the spread of the disease. • Although privacy concerns remain, when Bluetooth relies on Bluetooth Low Energy technology, all information is stored in the user's device (ie, a decentralized system), thus raising less privacy issues.

Technologies used for contact tracing apps	App examples, app name (country/region of origin)	Potential legal challenges
Self-reporting by users	• Epidemic Prevention and Health Information Code (China) • National Government Service Platform (China) • Self-quarantine safety protection (South Korea) • HealthLynked COVID-19 Tracker (United States) • Relief Central COVID-19 (United States) • PatientSphere for COVID-19 (United States) • Obvio-19 (United States) • How We Feel (United States) • COVID Safe Paths (United States) • CovidWatcher (New York City, United States) • Care19 (South Dakota, United States) • MyBellevue (Washington, United States) • Healthy Together (Utah, United States) • COVID Alert (Canada) • BC COVID-19 Support (British Columbia, Canada) • King's College London Covid-19 Symptom Reporting (United Kingdom) • NZ COVID Tracer (New Zealand)	• Self-reporting software usually requires users to upload their own personal data. The software analyzes the location information and personal health data of multiple users by means of big data. This involves methods for properly storing and using the personal information of users and other legal issues related to personal data privacy protection. • Under the absence of legal oversight, the self-reporting of health conditions by users can lead to the excessive collection of personal information by software. • Centralized data storage may lead to the improper access of information and excessive dissemination of users' personal information. • Some self-reporting software shares data with the authorities, which can make it possible for the government to access personal information. • There are no careful legal regulations to govern software and technology that collect data centrally for public health purposes in the context of pandemics. • Who will be responsible for any false information? Will users that provide false information be held accountable?

Different Jurisdictions Offer Different Ways to Protect Data

Existing domestic and regional data protection regulations are applicable for protecting users' personal information, but their solutions are not always compatible (Table 2). For instance, within the European Union, apps are subject to the requirements of the GDPR [15], and eventually to those of Directive 2002/58/EC [16]. Data processing must comply with basilar general principles (Article 5/a-e of the GDPR) [17], as follows: (1) data must be collected in a fair way, and data processing must be open to public scrutiny, as per the principle of lawfulness, fairness, and transparency; (2) the collected data are to be used for a specific and clear purpose, and changes in purpose are not allowed, as per the principle of purpose limitation; (3) when possible, anonymized or pseudoanonymized data shall be used, and measures must be taken to prevent the reidentification of the data subject, as per the first dimension of the principle of data minimization; (4) unnecessary information is not to be collected, as per the second dimension of the principle of data minimization; (5) data must be accurate, and any mistake shall be easily amended, as per the principle of accuracy; and (6) data are to be stored for a limited period of time and destroyed after that, as per the principle of storage limitation.

Table 2. Examples of domestic and regional data laws applicable to data collection and data sharing of contact tracing apps.

Countries or regions	Applicable data laws
The United States	Health Insurance Portability and Accountability Act
	Health Insurance Portability and Accountability Act Privacy Rule
	California Consumer Protection Act
The European Union	General Data Protection Regulation
	Directive 2002/58/EC
Canada	Health Information Act
	Freedom of Information and Protection of Privacy Act
China	Cyber Security Law of the People's Republic of China
	Personal Information Security Specifications
Australia	Privacy Amendment (Public Health Contact Information) Act

In the United States, although government-issued contact tracing apps exceed the scope of the Health Insurance Portability and Accountability Act (HIPAA) and HIPAA Privacy Rule, many contact tracing app developers claim that their apps will follow the requirements mandated by the HIPAA when collecting, storing, and using users' data [18]. Some California-based

companies further claim that their apps will also comply with the 2018 California Consumer Protection Act (CCPA) [19]. Even when only considering these 3 standards and disregarding all the remaining potentially applicable regulations, several different solutions can be found. Many of the rights guaranteed by the GDPR are unknown in US law, such as the right of data subjects to receive their personal data in a commonly used format and to transmit persona data to another data controller, the right to data portability, and the right to not to be subject to an adverse decision based solely on the application of artificial intelligence [20]. Moreover, the right to have one's data erased upon request, which is a basic right under the GDPR, is absent from the HIPAA. However, for apps governed by the CCPA, this right will apply, since the right was established in Civil Code § 1798.105 [21].

In Canada, contact tracing apps developed by public health authorities are subject to the data protection of both the Health Information Act (HIA) and the Freedom of Information and Protection of Privacy Act (FIPPA). For example, the ABTraceTogether contact tracing app developed in Alberta must adhere to the privacy obligations of the HIA and FIPPA [22]. Thus, the app only collects the user's nonidentifying information to trace the contacts of confirmed patients with COVID-19 [23]. In addition, based on the Public Health Act [24] and HIA [25], provincial health departments can track contacts during the COVID-19 pandemic and are able to use the information for health system management and planning, policy development, and public health emergency analysis. Unlike the GDPR, the FIPPA only applies to government institutions, and the protected personal information is narrower than the concept of personal data adopted by GDPR (eg, IP address and cookie data are not covered by the FIPPA) [26]. Moreover, provisions in the FIPPA [26] regarding data process and data subject rights are much less extensive than those speculated in the GDPR [27].

In China, the collection and processing of personal information garnered from COVID-19 tracking software is protected by laws and regulations that govern personal information protection and infectious disease prevention and control, such as the Cyber Security Law, the Personal Information Security Specifications, and the Regulations on Public Health Emergencies. Accordingly, when acquiring and using personal information, contact tracing apps shall obtain consent from the users in advance. The data processor shall also not intentionally disclose the personal information of confirmed or suspected patients with COVID-19 and their close contacts. Under the explicit authorization from users, disease prevention and control institutions and medical institutions can track high-risk populations based on legally obtained information. However, China's data protection laws and regulations are mainly focused on public security concerns, while few provisions are about personal data protection [28]. Data subject rights granted by these relevant laws are much narrower than those granted by the provisions in the GDPR.

It is important to note that data protection laws are generally lacking in low-income countries, and many have not promulgated privacy laws. According to the United Nations Conference on Trade and Development, 19% of low-income countries have no privacy legislation, and just 10% have developed a draft of their intended policies [29]. This privacy gap brings significant legal barriers to transnational data transfer and sharing.

Some Jurisdictions Allow for Transnational Data Transfer, While Others Do Not

Many domestic data protection regulations have already established conditions for transnational data transfer, and they are varied among jurisdictions. Consequently, the cross-border communication and transmission of epidemic contact tracing data will inevitably encounter legal challenges.

In China for example, data must undergo an internal security assessment that adheres to the regulations established by state network departments before data can be transferred outside the country. According to the Security Assessment Measures for Cross-Border Transfer of Personal Information (Draft for Comment) issued in 2019, network operators shall apply security assessments for the cross-border transfer of personal information to local cyberspace administrations at the provincial level before the personal information leaves China [30].

In the European Union, the GDPR imposes strict limitations on using a third party (ie, a country outside the European Economic Area or an international organization) to process data transfers (chapter V of the GDPR) [31], as this is likely to happen in international traveling. Transference can only take place if the law applicable by the third party "ensures an adequate level of protection", as confirmed by an adequacy decision taken by the European Commission (Article 45 of the GDPR) [31], or "if the controller or processor has provided appropriate safeguards" regarding data safety, the protection of the subject's rights, and the existence of adequate remedies (Article 46 of the GDPR) [31]. Apart from these 2 cases, only the specific and exceptional conditions laid down in Article 49 of the GDPR remain [31]. This specific set of scenarios for data transfer might hamper the operationality of non-European Union contact tracing apps. Contact tracing apps used in a country where data security protection does not meet the European Union's standards cannot obtain the infection risk information of travelers entering European Union member states. Hence, the destination country is not able to determine whether a traveler has already been infected before entering the country or has become infected after entry.

In some countries, data collected by contact tracing apps are not allowed to be transferred outside of the country. For example, in Australia, the Privacy Amendment (Public Health Contact Information) Act 2020 prohibits a person from disclosing data collected by contact tracing apps to another person outside of Australia [32]. The Australian COVIDSafe app also contains statements in the Terms of Service section that describes how all data collected by the government through the app are kept in Australia and cannot be transferred out of the country [33].

At the international level however, no mechanism enables the sharing of data essentials for pandemic control. Region-based contact tracing apps work as a technologic Tower of Babel, as they operate independently from one another and are not mutually recognized. The isolation of each app leaves a

monitoring gap for international travel, which leads to real challenges for global pandemic control.

Minimizing the International Health Risks Posed by This Technologic Tower of Babel

Possible Options for Destroying This Technologic Tower of Babel

In April 2020, both the European Parliament [34] and the European Commission [35] called on member states to work together to fight the pandemic. In June 2020, the European Data Protection Board released a statement regarding the impact of data protection on the interoperability of contact tracing apps within the European Union [36]. Several months later, in October 2020, a European Union-wide system for contact tracing app interoperability was launched, and the first group of contact tracing apps (ie, Germany's Corona-Warn-App, Ireland's COVID tracker app, and Italy's Immuni app) were linked to the system [37]. However, a broader consensus is required for the international community, since a common European approach will only solve the problem within the European Union.

The World Health Organization (WHO) has established ethical guidelines for directing member parties in the use of contact apps for the COVID-19 pandemic control [38]. Recently, the WHO Emergency Committee has also been actively working on the development of public health tools that can help member states deal with the pandemic-related risks involved with the gradual resumption of international travel [39]. These recommendations provide evidence-informed guidelines that member states can refer to when developing policies for pandemic control. To date, however, no international instrument has been established to address the conflicts of laws regarding data protection and privacy issues between member states. In this regard, we advocate that the global community should use the COVID-19 crisis as an opportunity for developing a mechanism that can facilitate the transnational sharing of data among countries during global public health emergencies. The biggest challenges will be reaching an agreement between countries and determining methods for enforcing such a mechanism.

Not long ago, the Global Outbreak and Response Network, which worked under the WHO, established a global data collection and sharing platform, Go.Data, which was used in South African countries for Ebola surveillance [40]. Ideally, a similar model that includes a large array of countries—potentially, all the affected ones—could be put in place for the COVID-19 pandemic. However, practical and legal impairments prevent this possibility. At most, we can aspire to create an open-access platform with aggregated data (ie, anonymized data) so that international travelers, institutions, and governments around the world can use the data to track the trends of infected persons and inquire about the risk of epidemic infection in a country or region.

Another bold proposal is a common contact tracing app that is accepted by all countries and made mandatory for international travelers, to help authorities track the flow of potentially infected international travelers. This international contact tracing app will fill the supervisory vacuum for international traveling, providing governments with real-time data to prevent the international spread of disease. Owing to the global citizenship obligation, international travelers should be required to install this tracing app [41]. This specific model for a contact tracing app would result from a common agreement. Governments should agree on a technological model for this app (Table 1). In any case, only the data necessary to perform the specific task assigned to this app should be collected. Such data should not include the biometric information of users, and the data collected should only be used for monitoring epidemic risk. As Google and Apple have worked together to promote contact tracing technology that allows communication between smartphones using iOS and Android systems via Bluetooth, this technologic progress may provide an option for minimizing international health risks. Based on Bluetooth and users' self-reports, this common contact tracing app would notify app users who have had close contact with a COVID-19–positive app user during international travel in the past 14 days.

Reality Check

Implementing any of these measures would not be easy in jurisdictions with rather stringent privacy laws. Each government has to accept a common data collection, processing, and transfer regime that is exclusively applicable to COVID-19 contact tracing apps as an exception to their respective national laws on privacy. This equates to a uniform privacy policy for all contact tracing app users (Figure 1).

Figure 1. Potential solutions (light blue box) to the technologic Tower of Babel and barriers to implementing these solutions (dark blue box).

Even with the leadership of the WHO and the menace of a global pandemic, the goal of fulfilling these measures is difficult to achieve. Governments will not easily give up the power to decide which contact tracing apps would be used in their territory and which laws should govern these apps (ie, their own laws). Even if they participate in such an international project, without global consensus on a legal framework for data collection, processing, and transfer, it is very unlikely that all countries can agree on a single model for contact tracing apps (Figure 1). Moreover, issues raised by international data transfer are another real challenge for achieving the interoperability of contact tracing apps. In particular, a recent decision made by the Court of Justice of the European Union (CJEU) indicates the legal challenges for transnational personal data transfer due to different data protection standards. In July 2020, the CJEU announced that the US-EU Privacy Shield, the personal data transfer arrangement between the United States and European Union, cannot provide sufficient data protection, and is therefore invalid [42]. This decision will dramatically impact personal data flows between the United States and European Union and will inevitably hamper the communication between contact tracing apps used in the 2 jurisdictions (Figure 1).

Conclusion

Contact tracing apps can be powerful mechanisms for handling the pandemic. However, their benefits will be lost if a patchwork system of noncommunicating apps becomes normal. If governments and app developers do not act cooperatively and strategically, the global community will be left to navigate dozens of apps operating under different protocols, data models, and legal rules, ultimately hampering the effort to control the COVID-19 pandemic.

Developing an international instrument that regulates data transfer between different countries' contact tracing apps is a challenging goal. The negotiation for such an instrument demands strong leadership and consistent efforts from the entire international community. Due to the reopening of international travel, we face the increased risk of future waves of COVID-19. Therefore, collaborative actions should be taken to minimize the risk of COVID-19 spread in international travel and transport. The international community should start working together to establish a framework that enables data sharing and data transfer among different contact tracing apps, thereby breaking this technologic Tower of Babel.

In spite of the existing difficulties, the goal of implementing these solutions is still worth pursuing, as it is for the common good for all human beings. The isolation of contact tracing apps

will not contain the international spread of infectious diseases. Instead, their isolation will result in the creation of a useless technologic Tower of Babel. The international community should be prepared for the next global pandemic crisis.

At present, it is crucial for all of us global citizens to collaborate with the common goal of establishing mechanisms that fix the current fragmentation of contact tracing systems and enable effective global epidemic surveillance. Only time will tell if national governments will primarily protect their power and regulations—which they feel is their sovereign right—under the price of paralyzing international traveling and jeopardizing our common safety. As it happened in biblical accounts, this technologic Tower of Babel will throw us apart.

Authors' Contributions

LD and VLR designed the study. MW collected and analyzed the data. LD and VLR wrote the manuscript.

Conflicts of Interest

None declared.

References

1. Chen S, Yang J, Yang W, Wang C, Bärnighausen T. COVID-19 control in China during mass population movements at New Year. Lancet 2020 Mar 07;395(10226):764-766 [FREE Full text] [doi: 10.1016/S0140-6736(20)30421-9] [Medline: 32105609]
2. 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]
3. Oliver N, Lepri B, Sterly H, Lambiotte R, Deletaille S, De Nadai M, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. Sci Adv 2020 Jun;6(23):eabc0764 [FREE Full text] [doi: 10.1126/sciadv.abc0764] [Medline: 32548274]
4. Li Z, Chen Q, Feng L, Rodewald L, Xia Y, Yu H, China CDC COVID-19 Emergency Response Strategy Team. Active case finding with case management: the key to tackling the COVID-19 pandemic. Lancet 2020 Jul 04;396(10243):63-70 [FREE Full text] [doi: 10.1016/S0140-6736(20)31278-2] [Medline: 32505220]
5. COVIDSafe app. Australian Government Department of Health. URL: https://www.health.gov.au/resources/apps-and-tools/covidsafe-app?gclid=EAIaIQobChMI6-ew3_7Q6wIVVz5gCh2VnwUoEAAYAiAAEgJ6Q_D_BwE [accessed 2020-10-30]
6. Download COVID Alert today. Government of Canada. URL: https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/covid-alert.html [accessed 2020-10-30]
7. Fernandez E. Privacy And Contact Tracing Apps - Google And Apple Debate With World Governments. Forbes. 2020 Apr 24. URL: https://www.forbes.com/sites/fernandezelizabeth/2020/04/24/privacy-and-contact-tracing-appsgoogle-and-apple-debate-with-world-governments/#1e837deb4518 [accessed 2020-10-30]
8. Jee C. Apple and Google have launched coronavirus exposure notifications without an app. MIT Technology Review. 2020 Sep 02. URL: https://www.technologyreview.com/2020/09/02/1007921/apple-and-google-have-launched-coronavirus-exposure-notifications-without-an-app/ [accessed 2020-10-30]
9. Mobile contact tracing apps in EU Member States. European Commission. URL: https://ec.europa.eu/info/live-work-travel-eu/health/coronavirus-response/travel-during-coronavirus-pandemic/mobile-contact-tracing-apps-eu-member-states_en [accessed 2020-10-30]
10. O'Neill P, 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. 2020 May 07. URL: https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker [accessed 2020-10-30]
11. Chidambaram S, Erridge S, Kinross J, Purkayastha S, PanSurg Collaborative. Observational study of UK mobile health apps for COVID-19. Lancet Digit Health 2020 Aug;2(8):e388-e390 [FREE Full text] [doi: 10.1016/S2589-7500(20)30144-8] [Medline: 32835196]
12. Kim HE. Coronavirus privacy: Are South Korea's alerts too revealing? BBC. 2020 Mar 05. URL: https://www.bbc.com/news/world-asia-51733145 [accessed 2020-10-30]
13. van der Eycken D. Europe's plan for contact tracing apps against COVID-19. CITIP blog. 2020 May 07. URL: https://www.law.kuleuven.be/citip/blog/europes-plan-for-contact-tracing-apps-against-covid-19/ [accessed 2020-10-30]
14. Li J, Guo X. Global deployment mappings and challenges of contact-tracing apps for COVID-19. SSRN. Preprint posted online on May 26, 2020. [doi: 10.2139/ssrn.3609516]
15. 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 2020-10-30]

16. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications). EUR-Lex. 2002 Jul 12. URL: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32002L0058 [accessed 2020-10-30]
17. Hoofnagle CJ, van der Sloot B, Borgesius FZ. The European Union general data protection regulation: what it is and what it means. Information & Communications Technology Law 2019 Feb 10;28(1):65-98 [FREE Full text] [doi: 10.1080/13600834.2019.1573501]
18. 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 May 28;7(1):lsaa034 [FREE Full text] [doi: 10.1093/jlb/lsaa034] [Medline: 32728470]
19. 1.81.5. California Consumer Privacy Act of 2018 [1798.100 - 1798.199]. California Legislative Information. 2018. URL: http://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=CIV&division=3.&title=1.81.5.&part=4.&chapter&article [accessed 2020-10-30]
20. Rothstein MA, Tovino SA. California Takes the Lead on Data Privacy Law. Hastings Cent Rep 2019 Sep;49(5):4-5. [doi: 10.1002/hast.1042] [Medline: 31581323]
21. Jehl L, Friel A. CCPA and GDPR Comparison Chart. Practical Law. URL: https://www.info-hisa.si/wp-content/uploads/2019/11/CCPA-GDPR-Chart.pdf [accessed 2020-11-21]
22. ABTraceTogether privacy statement. Government of Alberta. URL: https://www.alberta.ca/ab-trace-together-privacy.aspx [accessed 2020-11-21]
23. ABTraceTogether FAQ. Government of Alberta. URL: https://www.alberta.ca/ab-trace-together-faq.aspx [accessed 2020-11-21]
24. Public Health Act, RSA 2000, c P-37. CanLii. URL: https://www.canlii.org/en/ab/laws/stat/rsa-2000-c-p-37/168700/rsa-2000-c-p-37.html [accessed 2020-11-21]
25. Health Information Act, RSA 2000, c H-5. CanLii. URL: https://www.canlii.org/en/ab/laws/stat/rsa-2000-c-h-5/168705/rsa-2000-c-h-5.html [accessed 2020-11-21]
26. Freedom of Information and Protection of Privacy Act, RSO 1990, c F.31. CanLii. URL: https://www.canlii.org/en/on/laws/stat/rso-1990-c-f31/latest/rso-1990-c-f31.html [accessed 2020-11-23]
27. Chapter 3 Rights of the data subject. General Data Protection Regulation. URL: https://gdpr-info.eu/chapter-3/ [accessed 2020-11-21]
28. Feng Y. The future of China's personal data protection law: challenges and prospects. Asia Pacific Law Review 2019 Aug 05;27(1):62-82. [doi: 10.1080/10192557.2019.1646015]
29. Data Protection and Privacy Legislation Worldwide. UNCTAD. URL: https://unctad.org/page/data-protection-and-privacy-legislation-worldwide [accessed 2020-10-30]
30. Xinhua. China solicits opinions on security assessment of cross-border transfer of personal information. China Daily. 2019 Jun 13. URL: http://www.chinadaily.com.cn/a/201906/13/WS5d02240da3103dbf14328095.html [accessed 2020-11-21]
31. General Data Protection Regulation GDPR. General Data Protection Regulation. URL: https://gdpr-info.eu/ [accessed 2020-11-21]
32. Privacy Amendment (Public Health Contact Information) Act 2020. Australian Government Federal Register of Legislation. URL: https://www.legislation.gov.au/Details/C2020A00044 [accessed 2020-11-21]
33. The COVIDSafe Application Privacy Impact Assessment. Australian Government Department of Health. 2020 Apr 24. URL: https://www.health.gov.au/sites/default/files/documents/2020/04/covidsafe-application-privacy-impact-assessment-covidsafe-application-privacy-impact-assessment.pdf [accessed 2020-11-21]
34. EU coordinated action to combat the COVID-19 pandemic and its consequences. European Parliament resolution of 17 April 2020 on EU coordinated action to combat the COVID-19 pandemic and its consequences (2020/2616(RSP)). European Parliament. URL: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0054_EN.pdf [accessed 2020-10-30]
35. Commission Recommendation (EU) 2020/518 of 8 April 2020 on a common Union toolbox for the use of technology and data to combat and exit from the COVID-19 crisis, in particular concerning mobile applications and the use of anonymised mobility data. EUR-Lex. 2020 Apr 14. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32020H0518 [accessed 2020-10-30]
36. Statementon the data protection impact of the interoperability of contact tracing apps. European Data Protection Board. URL: https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_statementinteroperabilitycontacttracingapps_en_0.pdf [accessed 2020-10-30]
37. Coronavirus: EU interoperability gateway goes live, first contact tracing and warning apps linked to the system. European Commission. 2020 Oct 19. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1904 [accessed 2020-10-30]
38. WHO. Ethical considerations to guide the use of digital proximity tracking technologies for COVID-19 contact tracing. World Health Organization. 2020 May 28. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-Ethics_Contact_tracing_apps-2020.1 [accessed 2020-10-30]
39. Public health considerations while resuming international travel. World Health Organization. 2020 Jul 30. URL: https://www.who.int/news-room/articles-detail/public-health-considerations-while-resuming-international-travel [accessed 2020-10-30]

40. WHO's Data Tool improves Ebola Surveillance, Contact Tracing and Decision Making in Uganda. World Health Organization Uganda. 2019 Jul 26. URL: https://www.afro.who.int/news/whos-data-tool-improves-ebola-surveillance-contact-tracing-and-decision-making-uganda [accessed 2020-10-30]
41. Stoner L, Perry L, Wadsworth D, Stoner KR, Tarrant MA. Global citizenship is key to securing global health: the role of higher education. Prev Med 2014 Jul;64:126-128. [doi: 10.1016/j.ypmed.2014.05.006] [Medline: 24836370]
42. Data PCV. Judgment in Case C-311/18 Data Protection Commissioner v Facebook Ireland and Maximillian Schrems. Court of Justice of the European Union. 2020 Jul 16. URL: https://curia.europa.eu/jcms/upload/docs/application/pdf/2020-07/cp200091en.pdf [accessed 2020-10-30]

Abbreviations

API: application programming interface
CCPA: California Consumer Protection Act
CJEU: Court of Justice of the European Union
EPHIC: Epidemic Prevention and Health Information Code
FIPPA: Freedom of Information and Protection of Privacy Act
GDPR: General Data Protection Regulation
HIA: Health Information Act
HIPPA: Health Insurance Portability and Accountability Act
WHO: World Health Organization

Edited by L Buis; submitted 04.08.20; peer-reviewed by J Li, S Thiebes; comments to author 20.08.20; revised version received 07.09.20; accepted 03.11.20; published 27.11.20.

Please cite as:
Du L, Raposo VL, Wang M
COVID-19 Contact Tracing Apps: A Technologic Tower of Babel and the Gap for International Pandemic Control
JMIR Mhealth Uhealth 2020;8(11):e23194
URL: http://mhealth.jmir.org/2020/11/e23194/
doi:10.2196/23194
PMID:33156804

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org