JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 5 (2017), Issue 5 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Viewpoint

mHealth Assessment: Conceptualization of a Global Framework

Meghan Bradway[1,2*], MBA; Carme Carrion[3,4,5*], PhD; Bárbara Vallespin[6*], MBA; Omid Saadatfard[1*], MHI; Elisa Puigdomènech[7*], MPH; Mireia Espallargues[5,7*], PhD; Anna Kotzeva[7,8*], MPH

[1]Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway
[2]UiT The Arctic University of Norway, Department of Clinical Medicine, Tromsø, Norway
[3]Universitat Oberta de Catalunya (UOC), Catalonia, Spain
[4]TransLab Research Group, Departament de Ciències Mèdiques, Universitat de Girona (UdG), Catalonia, Spain
[5]Red de Investigación en Servicios Sanitarios en Enfermedades Crónicas (REDISSEC), Bilbao, Spain
[6]Mobile World Capital Barcelona, Barcelona, Spain
[7]Agència de Qualitat i Avaluació Sanitàries de Catalunya (AQuAS), Catalonia, Spain
[8]CIBER de Epidemiología y salud Pública (CIBERESP), Madrid, Spain
[*]all authors contributed equally

Corresponding Author:
Carme Carrion, PhD
Universitat Oberta de Catalunya (UOC)
Rambla del Poblenou, 156
Catalonia, 08018
Spain
Phone: 34 253 23 00
Fax: 34 417 64 95
Email: mcarrionr@uoc.edu

Abstract

Background: The mass availability and use of mobile health (mHealth) technologies offers the potential for these technologies to support or substitute medical advice. However, it is worrisome that most assessment initiatives are still not able to successfully evaluate all aspects of mHealth solutions. As a result, multiple strategies to assess mHealth solutions are being proposed by medical regulatory bodies and similar organizations.

Objective: We aim to offer a collective description of a universally applicable description of mHealth assessment initiatives, given their current and, as we see it, potential impact. In doing so, we recommend a common foundation for the development or update of assessment initiatives by addressing the multistakeholder issues that mHealth technology adds to the traditional medical environment.

Methods: Organized by the Mobile World Capital Barcelona Foundation, we represent a workgroup consisting of patient associations, developers, and health authority representatives, including medical practitioners, within Europe. Contributions from each group's diverse competencies has allowed us to create an overview of the complex yet similar approaches to mHealth evaluation that are being developed today, including common gaps in concepts and perspectives. In response, we summarize commonalities of existing initiatives and exemplify additional characteristics that we believe will strengthen and unify these efforts.

Results: As opposed to a universal standard or protocol in evaluating mHealth solutions, assessment frameworks should respect the needs and capacity of each medical system or country. Therefore, we expect that the medical system will specify the content, resources, and workflow of assessment protocols in order to ensure a sustainable plan for mHealth solutions within their respective countries.

Conclusions: A common framework for all mHealth initiatives around the world will be useful in order to assess whatever mHealth solution is desirable in different areas, adapting it to the specifics of each context, to bridge the gap between health authorities, patients, and mHealth developers. We aim to foster a more trusting and collaborative environment to safeguard the well-being of patients and citizens while encouraging innovation of technology and policy.

(JMIR Mhealth Uhealth 2017;5(5):e60) doi:10.2196/mhealth.7291

KEYWORDS

mhealth; evaluation; assessment; checklist; framework

Introduction

Mobile health (mHealth) is defined by the World Health Organization's (WHO) Global Observatory for eHealth as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices" [1]. Despite the lack of safety and quality validation of these technologies by medical regulatory bodies, individuals have adopted mHealth devices as self-management aids, while medical professionals are at a loss for how to relate to them [1]. Due to this consumer-based and rapid introduction within the world of patient health aids, mHealth solutions present unique and stakeholder-specific challenges to the medical environment. Patients, health care providers, administrators, authorities, and mHealth developers alike are operating without clear direction—from potentially improper use of mHealth apps by individuals, to medical systems' inability to react due to lack of technological and organizational support. Therefore, several questions arise: How should health authorities approach mHealth evaluation and certification? How can this be accomplished without stifling innovation? Should efforts to determine risks, benefits, and appropriate use be held on a global, country, or regional basis? How will policies and strategies be introduced to medical professionals and their practices? Which methods should be used to ensure that patients are properly informed on how to select and use mHealth solutions?

Today, there are many efforts underway to address these challenges [2-5]. However, these are happening in silos and are often specific to a single country or medical system. In this paper, we explore the progress and setbacks of the mHealth assessment environment and propose a collaborative and global approach to assessment, with the aim of applying the competencies of all stakeholder groups. We believe that by structuring our work in this field from a common foundation, assessment initiatives can and will foster a more trusting and collaborative environment to safeguard the well-being of citizens while encouraging innovation of technology and policy.

The purpose of this viewpoint paper is to be informative and provocative, focusing on the changes in mindset and actions that must occur—from the education and perspective of consumers, health assessment evaluations teams, and medical authorities—in the assessment of mHealth solutions. As such, we hope to elicit discussion on the content and methodologies of our suggested common framework for mHealth assessment.

The Potential of Mobile Technologies in the Medical Field

The potential of mHealth solutions lies in their ability to enable chronic diseases management and general wellness motivation [6-8]. By tracking an individual's health and lifestyle data, and by providing actionable feedback, these tools encourage self-management. With the subsequent promise of equipping health care providers with such detailed information about a patient's health status in relation to their daily habits, the medical system could not only better understand elusive compliance issues but also propose tailored solutions [9] for individual patients. If integrated and supported appropriately, these tools could improve treatment, empower patients, and foreseeably lower medical costs and streamline use of health care resources [10,11].

However, because these innovations are diverse and unstandardized, mHealth requires that regulatory bodies take a fresh look at evaluation methodologies for health aids. Today, medical actors are expected to adapt to the rapid technology turnover, ubiquity, and connectivity of mHealth by collecting, incorporating, and analyzing unprecedented amounts of personal health information seamlessly within the clinical environment. However, medical research and validation protocols, which providers rely on for guidance, are currently unable to adapt as quickly as such mHealth tools are being released. As a result, health care providers are left in the dark regarding how to relate to mHealth tools, as well as the impacts to their own professional responsibilities. If this goes un-checked, the medical system will continue to lag behind the needs of its patients.

The Current Situation

Just as in the proverbial scenario of the Band-Aid on the leaking dam, the mHealth environment has been met with a series of incomplete or issue-specific solutions to the evaluation of health apps, wearables and sensors. In fact, neither the Health Technology Assessment (HTA) agency [12] nor its collaborative networks, European Network for Health Technology Assessment (EUnetHTA) and HTA Network, have tackled mHealth solutions within their range of technology assessments. As patients continue to use these apps regardless of clinical support or guidance, health authorities and providers fall further behind, stuck in the traditional and hypercontrolled operations of the medical sector. With good intentions, organizations continuously attempt to address the distinct issues facing patients and clinicians. Promising attempts included the National Health Service's Health Apps Library [13], which aimed to involve clinicians in the review process, yet it is still under maintenance; the Organization for the Review of Care and Health Applications [14], which describes the purpose of their value and risk scores but not which features of an app lead to its value or risk; and PatientView's "The myhealthapps directory 2015-2016" [15], which presented a summary of consumer-generated app reviews. While these initiatives attempt to provide the public with digestible and relevant information, they simply add to the slew of disjointed and static solutions that are not able to address all stakeholder issues.

A Step in the Right Direction

Recently, governmental organizations and those representing end-user interests have acknowledged that flexible solutions and continuous communication are more appropriate for the mHealth environment than customary static reports. A

preparatory review of these evaluation initiatives illustrates the range of developed mHealth evaluation methodologies and adapted health assessment frameworks for use in mHealth. These initiatives focus mainly on the usability and clinical application of mHealth technologies, which are most commonly provided by individual user's commentaries, the organization's own evaluation team or, very rarely, a representative group of medical practitioners [16-19]. Some, including the United States Food and Drug Administration (FDA), choose to regulate only devices that fall under the definition of "medical device," enabling them to rely on existing frameworks for evaluation and implementation [20]. Others attempt to address a broader range of tools, with some focusing only on risk assessment [21,22] or even utilizing nontraditional resources (eg, novel technologies and social media outlets [23]) to evaluate mHealth solutions. Finally, most initiatives address only mHealth solutions that are fully operational within the market [24], with no solution for those under development [25].

Too Many Initiatives, Too Few Answers

Despite the diversity of approaches, recent reviews reveal worrisome results—a lack of conclusive or actionable evidence to suggest the ability of commercially available mHealth tools to effect behavioral changes or to manage chronic diseases, inpatient care, or health care delivery [26]. As a result, a primary concern and unfortunate reality is that this environment has allowed for the existence of misinformation regarding appropriate uses for these mHealth solutions. Common labels may even clearly categorize an app as "medical" yet include the warning in fine print that the app is "intended for entertainment only" [27]. Moreover, these evaluation efforts do not involve or inform all stakeholders [28] of relevant results, including potential risks associated with misleading health recommendations or software failures.

To compare the scope of structure and insights offered by today's mHealth evaluation initiatives to the desired coverage and impact that we believe is possible, we have completed a comparative summary of representative evaluation efforts (see Multimedia Appendix 1 [29-33]). These characteristics are in addition to the assumed basic coverage of usability and security. Such a lack of flexible and inclusive evaluation options has resulted in a lack of empirically demonstrated understanding of the benefits and risks that mHealth solutions provide to care delivery and health management. There is clearly a need to create an environment that explores additional options for mHealth evaluation. For example, individuals outside the medical realm have not traditionally had a majority voice in the assessment of health aids. Yet today, individual app users are becoming more knowledgeable about the daily benefits of mHealth and technology developers are exploring the potential of ever-present self-management systems. Therefore, we should consider these groups as underutilized resources—we have an opportunity to recognize and enthusiastically encourage the insight that these individuals can now contribute to mHealth evaluation efforts.

The Unsung Role of Developers in Assessment

Developers are the key force driving the need for certification due to the astounding rate at which they are producing mHealth solutions [34]. Therefore, those who are developing mHealth solutions should be encouraged to seek assessment for their own technologies through engagement in the evaluation process. Additionally, they should be provided with guidance and educational resources, including explanation of the concepts and benefits of a quality assured and reliable product go hand-in-hand with growth in the market and success, as the result of consumer engagement, loyalty, and trust [35,36]. By creating a mutually beneficial situation for developers and the medical system, evaluation frameworks can facilitate an environment that is transparent, trustworthy, and safe for users. It must also be noted that we must achieve a balance between development oversight and creative freedom. Too much regulation, or an unnecessarily lengthy process, would paralyze certification and inhibit adoption of mHealth solutions, which is evident from responses to the US FDA's complex and unclear rulings [37]. Therefore, involving developers in the improvement and operation of any mHealth assessment initiatives will ensure the safety of their creativity and the competitive, open health-app market.

The Need to Work From a Common Foundation

While the resources, evidence, and financial support for clinical implementation undoubtedly vary between countries and medical systems, there are shared challenges and themes for assessment. By acknowledging and addressing these common needs, we as a health care community can work in parallel as opposed to reinventing the wheel. Furthermore, these needs cannot hope to be addressed by a single organization or single perspective (eg, safety and privacy vs usability and consumer cost). The needs presented by the mHealth assessment arena can be met only through participation of all stakeholders, hand-in-hand with regulation and legislation, constantly adapting to updates in standards and the capacity of technology.

To facilitate the success of this collaborative approach, we must simultaneously change the paradigm of evaluation to more appropriately relate to the particularities of mHealth. As such, it must be continuous and iterative, while at the same time provide timely conclusions and actionable recommendations for improvement and implementation.

Acknowledging and Merging Universal Needs: Approaching the Solution Together

Organized by the Mobile World Capital Barcelona Foundation, the co-authors of this paper represent a workgroup of several international partners with distinct yet complementary backgrounds and competencies related to mHealth, including patient groups, government, and health authorities, as well as representatives from health research, care providers, and technology developer groups. Exploration of the common needs and themes of mHealth has resulted in a summative list of concepts that we believe are universal to all assessment initiatives. We welcome interested parties and future partners to contribute their competencies and applicable solutions toward

the future development of common and foundational guidelines for building mHealth assessment frameworks. Not only are newly validated tools needed in order to improve the quality of assessment steps, but also the perspectives of a diverse group of representatives in all affected fields. The following sections outline the major universal concepts and approaches that can inform proposed and existing evaluation initiatives.

Assigning the Evaluation Team

The concept of specialization is a key component in any effort to impact a multistakeholder system. A single organization cannot be expected to accomplish the diversity of tasks and successfully address the challenges of mHealth evaluation (eg, results of safety, usability, and health change assessments). Therefore, representative organizations should be involved in tasks associated with their competencies, thus providing even distribution of responsibilities as well as relevant input.

In order to achieve a comprehensive process, the evaluation team should possess a broad scope of perspectives that are representative of the following stakeholder groups: patients or patient organizations; commercial and research-based mHealth developers; health care providers, medical professionals, and system administrators; insurance or other reimbursement bodies; and authorities within governmental health and medical system organizations.

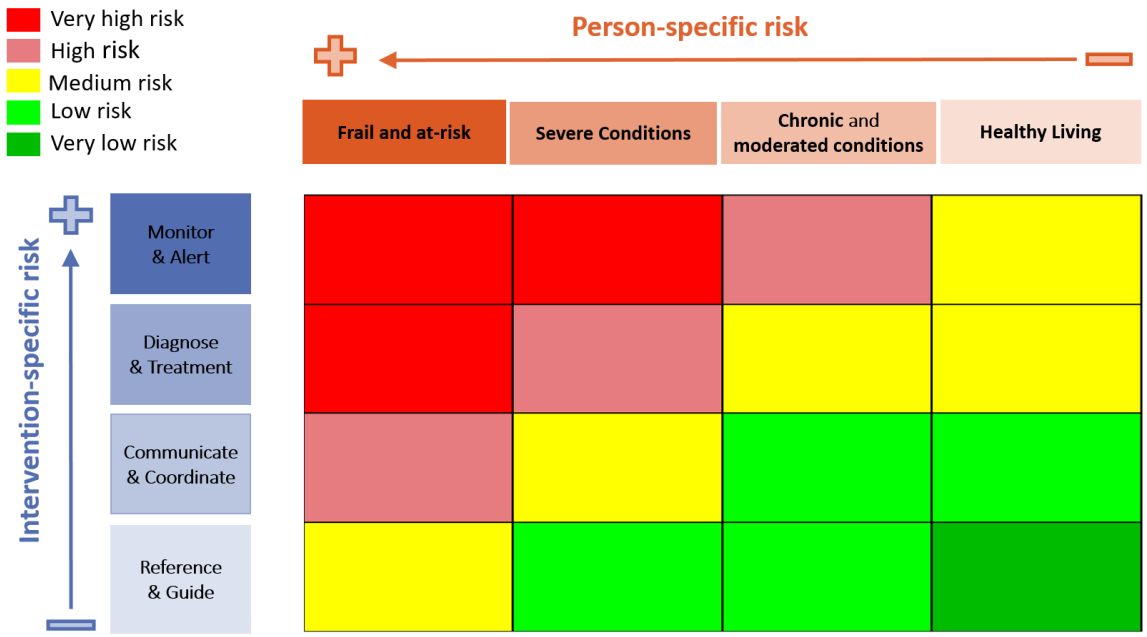
It is assumed that the composition of this team, and specific roles of stakeholder representatives, will vary based on the decision of each country or region that chooses to adopt and adapt this approach to their own respective medical system.

Pre-Assessment of mHealth Solutions

The pre-assessment phase is meant to classify any mHealth app and can be based on, for example, the following criteria: (1) Risk classification matrix: combining intervention type and patient type (see Figure 1 [38] for an example); (2) Users: patients, medical professionals, and/or informal caregivers; and (3) Integration: stand-alone, partially integrated, or fully integrated.

In order to address mHealth solutions within any stage of development, pre-assessment guides should be distributed to developers who are still in the process of designing health solutions as well as for members of evaluation teams within any medical system or country.

Figure 1. Risk assessment matrix.



Checklists

Together with the pre-assessment phase, checklists can act as preparative resources for the evaluation teams to ensure that all relevant information is provided before resources and time are spent assessing an mHealth solution, as well as subsequent public and medical integration. Once pre-assessment has been established, checklists can be directed toward each contextual category of health apps. The checklist will vary from country to country and be provided to the evaluation team. To illustrate this point, a checklist should ensure that documentation related to the following categories is provided for each mHealth solution:

- designation of mHealth solutions by intended use, for example, reference guides (eg, for nutrition or weight control), monitoring devices (eg, for blood sugars or blood pressure), or other types of solutions within a matrix
- level of development
- security and privacy
- interoperability standards
- usability
- functionalities and content

To the best of our knowledge, there are few functional and representative checklists (eg, Catalonia, Andalucía, and WHO's mHealth Technical Evidence Review Group [29,39,40]). Most

others are not fully available, are under development (eg, European Commission mHealth assessment working group [5]), or are focused only on mobile medical devices (eg, Future Internet-STAR checklist model [41]). These checklists propose categories related to functionality including usability, technology, security, content design and pertinence, and services. They also propose more domains related to infrastructure, including intervention delivery, accessibility toward individuals (ie, barriers or facilitators to the adoption of the intervention among study participants), cost assessment, adoption inputs/program entry, limitations for delivery at scale, contextual adaptability, and replicability. Many of these approaches involve classification of information into levels, from mandatory to not applicable, based on the app's level of risk, thereby providing flexibility to assess not only medical but also nonmedical devices.

Again, each stakeholder representative within evaluation teams should be assigned to the topics they are most capable of assessing. However, not all of these available checklists are accessible or usable by all stakeholders. As previously mentioned, developers are a key stakeholder in the success of assessment initiatives. Therefore, all developers should be able to provide the checklist information themselves. An alternative and more time-consuming option is that facilitators be put in place to gather this information from multiple sources in order to answer evaluation questions. This will affect the level of time and funded investment needed, which supports the involvement and support of developer groups in the design of checklists.

mHealth Evaluation Aspects and Methods

Assessment initiatives must be equally focused on summative (ie, during or post implementation) as well as formative (throughout the development life cycle) evaluation. This will ensure that not only common software development life-cycle stages but also end-user needs and input are incorporated, thus giving any initiative practical dimension. We expect that in doing so, this will promote adoption by the app development community as well as those representing the interests of the medical community.

The categories of information provided by checklists can be evaluated with the common evaluation perspectives of technical readiness and maturity, risks, benefits, and resources needed. The limitations and specifics of what is involved in these evaluation domains should correspond to the purpose of the evaluation (ie, assessing a device as an educational or medical tool) and should vary depending on the level of interoperability and intended use (eg, disease self-management vs activity tracking) of the mHealth solution (eg, stand-alone app vs integrated medical device).

Each of these four domains and their subdomains should be defined to streamline evaluation efforts and organize stakeholder participation based on their respective competencies (see Multimedia Appendix 2 for an example). Ideally, the chosen set of domains and subdomains should address the following needs:

- Determine the appropriate use of each mHealth solution (ie, as a medical device or a health and wellness tool, based on the target and breadth of functionalities as well as status of interoperability and safety standards)
- Develop expedited and conclusive methods to evaluate the effect(s) that an mHealth solution has on respective clinical outcomes and/or patient lifestyle habits, based on its appropriate and intended use
- Assess risk related to (1) patients and their caregivers in relation to personal data security, self-management decision making, and disease understanding; (2) clinicians, including liability to their practice and a greater trust of and reliance on patient-gathered data; and (3) overall health care organizations and systems, including financial impact and liabilities
- Inform stakeholders of relevant results through respective and accessible platforms.

Quantitative or qualitative methodologies may be employed through formal evaluation studies (eg, clinical trials) to assess each of these. The evaluation studies should be conducted based on the regulations, practices, and implementation strategies of individual countries or medical systems in which such studies are completed. However, to our knowledge, most of the subdomains still lack validated, standardized, or descriptive approaches and methods for evaluation. Therefore, alternative or unconventional inquiry methods should be considered including involvement of medical education programs as well as clinical and commercial research organizations, all of which use complementary methods such as online platforms, market analysis, and user-involved workshops. Partnership with organizations that are able to develop such novel methods should be sought both within the international realm of health sciences and then adapted to the needs and socioeconomic factors present within each unique medical system.

Conclusions

There is a clear need for defining a standard assessment framework for mHealth technologies that will help separate the wheat from the chaff and identify those solutions that may provide added value to patients and the health care system. By accomplishing this set of methodologies and approaches to evaluation and also the methods, perspectives, and resources used to accomplish these tasks, evaluation would facilitate more educated and informed decision making regarding the choice and use of mHealth solutions—from patients to medical practitioners to the health authorities that are charged with maintaining the foundation of each medical service.

The insights and suggestions provided in this paper are intended for the groups that are completing evaluations, developing apps, and creating health policy and infrastructural support for mHealth implementation. We propose that all organizations and individuals who share a similar passion for a coordinated effort towards a more rigorous and comprehensive evaluation of mHealth technologies join forces to form a virtual community of practice [42]. Discussing the merits and shortcomings of the proposed approach in this paper and its utility in real-world scenarios can be a starting point for such community. We hope that a free-flowing format will foster creative ways to solve the problems put forward in this position paper. A social

media based discussion group is a feasible starting point for this virtual community of practice. The authors invite the experts in evaluation of health information systems to provide their opinion as to how such a virtual community should form and collaborate.

Authors' Contributions

This paper was prepared by all the authors in their capacity. Authors EP and CC can be contacted by those interested in this discussion.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Presence of desirable characteristics within existing mHealth assessment initiatives.

[PDF File (Adobe PDF File), 82KB - mhealth_v5i5e60_app1.pdf]

Multimedia Appendix 2

Example set of domains and subdomains for an mHealth assessment framework.

[PDF File (Adobe PDF File), 327KB - mhealth_v5i5e60_app2.pdf]

References

1. World Health Organization. Frequently asked questions on Global Task Force on digital health for TB and its work. URL: http://www.who.int/tb/areas-of-work/digital-health/faq/en/ [accessed 2017-02-27] [WebCite Cache ID 6oaKq7nal]
2. Chant L, Hoogewerf L, Lewis T, Sathanandam S, Thimbleby H, Wyatt J. Rapid literature review for the National Information Board work stream 1.2.: Royal College of Physicians; 2015 Aug 07. URL: http://harold.thimbleby.net/NICE/NIB%20literature%20review%20final%20report%20v096-harold%20edits.pdf [accessed 2016-10-19] [WebCite Cache ID 6lNHQ6mM0]
3. World Health Organization, UN Foundation, Johns Hopkins University Global mHealth Initiative. mHealth MAPS toolkit: mHealth Assessment and Planning for Scale. 2015. URL: http://who.int/life-course/publications/mhealth-toolkit/en/ [accessed 2016-10-19] [WebCite Cache ID 6lNKtvDjE]
4. Haugli H. Høring: Forslag til selvdeklareringsordning for mobile helseapplikasjoner. Oslo: Direktoratet for ehelse; 2016 Aug 30. URL: https://ehelse.no/Documents/E-helsekunnskap/H%c3%b8ringsnotat-%20Forslag%20selvdeklareringsordning%20mobile%20helseapplikasjoner.pdf [accessed 2016-10-19] [WebCite Cache ID 6lNSgPz24]
5. European Commission. 2016 Jan 15. Working group on mHealth assessment guidelines URL: http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetail&groupID=3390 [accessed 2016-10-19] [WebCite Cache ID 6lNSxGJGO]
6. Talboom-Kamp E, Verdijk N, Harmans L, Numans M, Chavannes N. An eHealth Platform to Manage Chronic Disease in Primary Care: An Innovative Approach. Interact J Med Res 2016 Feb 09;5(1) [FREE Full text] [doi: 10.2196/ijmr.4217] [Medline: 26860333]
7. 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) [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
8. Garg S, Garg D, Turin TC, Chowdhury MF. Web-Based Interventions for Chronic Back Pain: A Systematic Review. J Med Internet Res 2016 Jul 26;18(7):e139 [FREE Full text] [doi: 10.2196/jmir.4932] [Medline: 27460413]
9. Azar KM, Koliwad S, Poon T, Xiao L, Lv N, Griggs R, et al. The Electronic CardioMetabolic Program (eCMP) for Patients With Cardiometabolic Risk: A Randomized Controlled Trial. J Med Internet Res 2016 May 27;18(5):e134 [FREE Full text] [doi: 10.2196/jmir.5143] [Medline: 27234480]
10. Hung CS, Yu JY, Lin YH, Chen YH, Huang CC, Lee JK, et al. Mortality Benefit of a Fourth-Generation Synchronous Telehealth Program for the Management of Chronic Cardiovascular Disease: A Longitudinal Study. J Med Internet Res 2016 May 13;18(5) [FREE Full text] [doi: 10.2196/jmir.5718] [Medline: 27177497]
11. Badawy SM, Kuhns LM. Economic Evaluation of Text-Messaging and Smartphone-Based Interventions to Improve Medication Adherence in Adolescents with Chronic Health Conditions: A Systematic Review. JMIR Mhealth Uhealth 2016 Oct 25;4(4) [FREE Full text] [doi: 10.2196/mhealth.6425] [Medline: 27780795]
12. EUnetHTA. EUnetHTA Joint Action 2 on HTA 2012-2015: 3-year Work Plan. 2013 May 01. URL: http://www.eunethta.eu/sites/5026.fedimbo.belgium.be/files/EUnetHTA%20JA2%203-year%20Work%20Plan%20%282013%29.pdf [accessed 2017-03-14] [WebCite Cache ID 6oxEcGGBe]

13. Hall K. The Register.: Situation Publishing; 2015 Jul 09. Privacy campaigners question credibility of NHS 'endorsed' apps URL: https://www.theregister.co.uk/2015/07/09/credibility_of_nhs_endorsed_apps_in_question/ [accessed 2017-03-14] [WebCite Cache ID 6oxExcXi4]
14. Organisation for the Review of Care and Health Applications. WHo are ORCHA?.: ORCHA; 2017. URL: https://www.orcha.co.uk/ [accessed 2017-03-14] [WebCite Cache ID 6oxFMsRps]
15. My Health Apps: tried and tested by people like you.: Patient View; 2013. URL: http://www.patient-view.com/bull-myhealthappsnet.html [accessed 2017-04-26] [WebCite Cache ID 6q0efjtSV]
16. Misra S, Husain I. iMedicalApps.: iMedicalApps; 2015 Dec 21. The Best Medical Apps Released in 2015 for iPhone & Android URL: http://www.imedicalapps.com/2015/12/best-medical-apps-iphone-android/ [accessed 2016-10-20] [WebCite Cache ID 6lOXV6kr5]
17. UMass Center for mHealth and Social Media. Massachusetts: University of Massachusetts Medical School Prevention and Behavioral Medicine App Evaluation Service URL: http://www.umassmed.edu/mhealth/app-evaluation-service/ [accessed 2016-10-20] [WebCite Cache ID 6lOYJEi0O]
18. Singh K, Drouin K, Newmark LP, Rozenblum R, Lee J, Landman A, et al. Developing a Framework for Evaluating the Patient Engagement, Quality, and Safety of Mobile Health Applications. Issue Brief (Commonwealth Fund) 2016 Feb;5:1-11 [FREE Full text] [Medline: 26934758]
19. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. Transl Behav Med 2014 Dec;4(4):363-371 [FREE Full text] [doi: 10.1007/s13142-014-0293-9] [Medline: 25584085]
20. US Food and Drug Administration. Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff. 2015 Feb 09. URL: https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm263366.pdf [accessed 2017-02-27] [WebCite Cache ID 6oaMDOBNW]
21. Wyatt JC, Thimbleby H, Rastall P, Hoogewerf J, Wooldridge D, Williams J. What makes a good clinical app? Introducing the RCP Health Informatics Unit checklist. Clin Med (Lond) 2015 Dec;15(6):519-521. [doi: 10.7861/clinmedicine.15-6-519] [Medline: 26621937]
22. Lewis TL, Wyatt JC. mHealth and mobile medical Apps: a framework to assess risk and promote safer use. J Med Internet Res 2014 Sep 15;16(9):e210 [FREE Full text] [doi: 10.2196/jmir.3133] [Medline: 25223398]
23. Facebook. London: PatientView; 2013 Nov 01. myhealthapps.net URL: https://www.facebook.com/myhealthapps [accessed 2016-10-20] [WebCite Cache ID 6lOe94RYv]
24. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
25. Royal Academy of Engineering. Health apps: regulation and quality control.: Academy of Medical Sciences; 2015. URL: http://www.raeng.org.uk/publications/reports/health-apps-regulation-and-quality-control [accessed 2016-10-20] [WebCite Cache ID 6lOfB7ly9]
26. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283). [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
27. Krisch J. Questioning the Value of Health Apps.: The New York Times; 2015 Mar 16. URL: http://well.blogs.nytimes.com/2015/03/16/health-apps-provide-pictures-if-not-proof-of-health/?_r=1 [accessed 2016-10-20] [WebCite Cache ID 6lOhPnGTB]
28. MyHealthApps. Health apps: where do they make sense?.: Patient View; 2014 Mar. URL: http://www.ehfg.org/blog/wp-content/uploads/2014/03/Health-apps-patients-standards-and-policies-white-paper-March-2014.pdf [accessed 2016-10-20] [WebCite Cache ID 6lOfcMpag]
29. Agencia de Calidad Sanitaria de Andalucía. Safety and Quality Strategy in Mobile Health Apps. 2012. AppSaludable Quality Seal URL: http://www.calidadappsalud.com/distintivo-appsaludable/ [accessed 2016-10-20] [WebCite Cache ID 6lOr20PU7]
30. Connect DG. Digital Single Market: Digital Economy & Society.: European Commission; 2016 May 04. Open stakeholder meeting on mHealth assessment guidelines: Presentations and survey URL: https://ec.europa.eu/digital-single-market/en/news/open-stakeholder-meeting-mhealth-assessment-guidelines-presentations-and-survey [accessed 2016-10-20] [WebCite Cache ID 6lOqKGoAa]
31. TicSalut: Technologia, innovacio i salut.: TicSalut - Generalitat de Catalunya; 2013. Standards and interoperability URL: http://www.ticsalut.cat/estandards/en_index/ [accessed 2016-10-20] [WebCite Cache ID 6lOqlgn8K]
32. Future Internet: Social Technological Alignment in Healthcare. 2014. Legal compliance for developers using FI-STAR eHealth platform: Training Materials URL: https://www.fi-star.eu/fileadmin/images/use_cases/Microsoft_PowerPoint_-_Introduction_into_applicable_legal_requirements.pdf [accessed 2016-10-20] [WebCite Cache ID 6lOrFqGA3]
33. Future Internet: Social and Technological Alignment in Healthcare. 2013. Basic validation methodology and quality control indicators URL: https://bscw.fi-star.eu/pub/bscw.cgi/d44362/D6.

1-Basic%20validation%20methodology%20and%20Quality%20control%20indicators.pdf [accessed 2016-10-19] [WebCite Cache ID 6lOrZWxE6]
34. Krohn R, Metcalf D, editors. mHealth Innovation: Best Practices From The Mobile Frontier. Chicago: Healthcare Information and Management Systems Society; 2014.
35. Batchelor C. United Kingdom Regulators Network. London: UK Regulators Network; 2014 Dec 17. Consumer engagement and switching URL: http://www.ukrn.org.uk/wp-content/uploads/2016/07/20141217ConsumerEngSwitch.pdf [accessed 2016-10-20] [WebCite Cache ID 6lOoMX9tr]
36. Halim P, Swasto B, Hamid D, Firdaus MR. European Journal of Business and Management. 2014. The Influence of Product Quality, Brand Image, and Quality of Service to Customer Trust and Implication on Customer Loyalty (Survey on Customer Brand Sharp Electronics Product at the South Kalimantan Province) URL: http://iiste.org/Journals/index.php/EJBM/article/viewFile/16089/16236 [accessed 2017-02-27] [WebCite Cache ID 6oaMbd2Rg]
37. Cortez NG, Cohen IG, Kesselheim AS. FDA regulation of mobile health technologies. N Engl J Med 2014 Jul 24;371(4):372-379 [FREE Full text] [doi: 10.1056/NEJMhle1403384] [Medline: 25054722]
38. Vallespin B, Cornet J, Kotzeva A. Ensuring evidence-based safe and effective mHealth applications. Evidence-Based Health Informatics 2016;222:248-261 [FREE Full text] [doi: 10.3233/978-1-61499-635-4-248] [Medline: 27198107]
39. TicSalut: Technologia, innovacio i salut.: TicSalut - Generalitat de Catalunya; 2013. Observatory URL: http://www.ticsalut.cat/observatori/en_index/ [accessed 2016-10-20] [WebCite Cache ID 6lOqlgn8K]
40. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar;352 [FREE Full text] [doi: 10.1136/bmj.i1174] [Medline: 26988021]
41. FI-STAR: Future Internet Social and Technological Alignment in Healthcare.: FI-STAR Consortium; 2014. Information for application developers using FI-STAR eHealth platform URL: https://www.fi-star.eu/publications/checklist-for-app-developers.html [accessed 2016-10-20] [WebCite Cache ID 6lP2FTr6n]
42. Ford J, Korjonen H, Keswani A, Hughes E. Virtual communities of practice: can they support the prevention agenda in public health? Online J Public Health Inform 2015 Jul;7(2) [FREE Full text] [doi: 10.5210/ojphi.v7i2.6031] [Medline: 26284150]

Abbreviations

EUnetHTA: European Network for Health Technology Assessment
FDA: Food and Drug Administration
FI-STAR: Future Internet Social and Technological Alignment Research
HTA: Health Technology Assessment Agency
mHealth: mobile health
WHO: World Health Organization

Edited by M Focsa; submitted 10.01.17; peer-reviewed by I Montagni, B White; comments to author 31.01.17; revised version received 14.03.17; accepted 14.03.17; published 02.05.17.

Please cite as:
Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, Kotzeva A
mHealth Assessment: Conceptualization of a Global Framework
JMIR Mhealth Uhealth 2017;5(5):e60
URL: http://mhealth.jmir.org/2017/5/e60/
doi:10.2196/mhealth.7291
PMID:28465282

Review

Smart Devices for Older Adults Managing Chronic Disease: A Scoping Review

Ben YB Kim[1], MHI; Joon Lee[1], PhD

Health Data Science Lab, School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Joon Lee, PhD
Health Data Science Lab
School of Public Health and Health Systems
University of Waterloo
Lyle Hallman North, 3rd Floor
200 University Avenue W
Waterloo, ON, N2L 3G1
Canada
Phone: 1 519 888 4567 ext 31567
Fax: 1 519 746 6776
Email: joon.lee@uwaterloo.ca

Abstract

Background: The emergence of smartphones and tablets featuring vastly advancing functionalities (eg, sensors, computing power, interactivity) has transformed the way mHealth interventions support chronic disease management for older adults. Baby boomers have begun to widely adopt smart devices and have expressed their desire to incorporate technologies into their chronic care. Although smart devices are actively used in research, little is known about the extent, characteristics, and range of smart device-based interventions.

Objective: We conducted a scoping review to (1) understand the nature, extent, and range of smart device-based research activities, (2) identify the limitations of the current research and knowledge gap, and (3) recommend future research directions.

Methods: We used the Arksey and O'Malley framework to conduct a scoping review. We identified relevant studies from MEDLINE, Embase, CINAHL, and Web of Science databases using search terms related to mobile health, chronic disease, and older adults. Selected studies used smart devices, sampled older adults, and were published in 2010 or after. The exclusion criteria were sole reliance on text messaging (short message service, SMS) or interactive voice response, validation of an electronic version of a questionnaire, postoperative monitoring, and evaluation of usability. We reviewed references. We charted quantitative data and analyzed qualitative studies using thematic synthesis. To collate and summarize the data, we used the chronic care model.

Results: A total of 51 articles met the eligibility criteria. Research activity increased steeply in 2014 (17/51, 33%) and preexperimental design predominated (16/50, 32%). Diabetes (16/46, 35%) and heart failure management (9/46, 20%) were most frequently studied. We identified diversity and heterogeneity in the collection of biometrics and patient-reported outcome measures within and between chronic diseases. Across studies, we found 8 self-management supporting strategies and 4 distinct communication channels for supporting the decision-making process. In particular, self-monitoring (38/40, 95%), automated feedback (15/40, 38%), and patient education (13/40, 38%) were commonly used as self-management support strategies. Of the 23 studies that implemented decision support strategies, clinical decision making was delegated to patients in 10 studies (43%). The impact on patient outcomes was consistent with studies that used cellular phones. Patients with heart failure and asthma reported improved quality of life. Qualitative analysis yielded 2 themes of facilitating technology adoption for older adults and 3 themes of barriers.

Conclusions: Limitations of current research included a lack of gerontological focus, dominance of preexperimental design, narrow research scope, inadequate support for participants, and insufficient evidence for clinical outcome. Recommendations for future research include generating evidence for smart device-based programs, using patient-generated data for advanced data mining techniques, validating patient decision support systems, and expanding mHealth practice through innovative technologies.

(JMIR Mhealth Uhealth 2017;5(5):e69) doi:10.2196/mhealth.7141

KEYWORDS

mobile health; mHealth; smartphone; mobile phone; tablet; older adults; seniors; chronic disease; chronic disease management; scoping review

Introduction

The advent of the cellular phone and its wide adoption by the general public intrigued many medical and health researchers. mHealth—the delivery of health care services and information via mobile technologies—has provided unique benefits to fulfill these needs because they are portable, affordable, widely available, and widely adopted [1-3]. mHealth research with cellular phones for chronic disease management has been extensive, and many systematic reviews have concluded that this technology has had a positive but limited impact on clinical outcomes, including improved hemoglobin A_{1c} (HbA_{1c}) control for people with diabetes [2,4,5] and tighter blood pressure control among hypertensive patients [6-8].

In the last few years, the boundaries of mHealth have been further expanded, driven by the rapid advancements in technological capabilities through the integration of an array of sensors (eg, accelerometer, pulse oximeter), multitouch screens, and higher computing power [9]. Most importantly, smartphones and tablets run apps that provide broader functionalities beyond their core features [10]. Owing to these advancements, smartphones and tablets have been used differently, most notably, through the collection of patient data with an array of sensors, the use of decision-making algorithms, and the presentation of visually augmented and interactive longitudinal data [11,12]. Such vastly different utility warrants a distinction between the traditional mHealth interventions based on cellular phones and modern smart device-based interventions. We define *smart device-based mHealth interventions* as mHealth interventions that use a smartphone, tablet, or peripheral device and their unique features, including high computing power, advanced functions using apps, use of sensors, fast network speed, and the interactive multitouch screen, to deliver health care services.

Smart devices have inherited the characteristics of cellular phones, and they are portable, pervasive, and becoming more and more affordable. In early 2016, the adoption rates of smartphones and tablets by older adults aged 60 to 69 years were 46% and 41%, respectively [13]. Older adults also have expanded the way they use smartphones and tablets, with health information seeking being the second most frequently executed task besides making phone calls [14]. With the growing older population, it is expected that more and more older adults will be interested in incorporating these smart devices into their chronic disease management.

As smart devices started to be widely adopted in the early 2010s, a reflective increase in mHealth research activities beginning in 2010 was noted in previous studies [11,12,15]. The randomized controlled trial (RCT) was identified as the most common research design, followed by descriptive and feasibility studies [11]. One study reported texting (short message service, SMS) as the most frequently used intervention tool, followed by apps on smartphones, noting its significant and growing role [11]. Two systematic studies reported overall positive patient outcomes from mHealth interventions but did not specify which clinical measures had the most benefit [11,15]. Another scoping review examined the design, development, and evaluation of mHealth systems for community-living older adults [12]. This study identified significant strategies for successful adoption among older adults, including user-centered design, collaborative team approaches, change management tactics for improving organizational and systems readiness, and formative and summative evaluation for improved user interfaces [12]. However, this study did not assess the characteristics and nature of smart device-based interventions and their impact on patient outcome. Overall, previous research had a vague distinction between the traditional and smart device-based interventions and lacked focus on the geriatric population.

To accommodate the broader functionalities and wide adoption of smart devices that can potentially better support chronic disease management for older adults, there is a need to synthesize research activities and evidence regarding smart device-based interventions. This scoping review explored this knowledge gap by mapping the extent and nature of the available literature, and will provide future directions for smart device-based interventions for older adults managing chronic disease.

Methods

Study Design

We chose to conduct a scoping review because the topic of smart device-based mHealth has emerged across a wide range of disciplines and it includes diverse chronic conditions. In particular, we used the Arksey and O'Malley framework for a scoping review [16] and the revisions made by Levac et al [17]. We also used thematic synthesis of qualitative studies to contextualize the findings of the quantitative studies.

Framework Stage 1: Identifying the Research Question

The first stage began with the previously described literature review, followed by the identification of the knowledge gap. To comprehensively map the extent and nature of smart device-based mHealth research activities for chronic disease management for older adults, this scoping review set out to answer 4 research objectives: (1) to understand the range and methods of biometric measurements and patient-reported outcome measures (PROMs) in smart device-based mHealth interventions, (2) to describe the types of self-management support employed in smart device-based mHealth interventions, (3) to describe the impact of smart device-based mHealth interventions on clinical outcomes, and (4) to thematically synthesize older adults' experience with smart device-based mHealth systems.

Framework Stage 2: Identifying Relevant Studies

We searched 4 databases—MEDLINE, Embase, CINAHL, and Web of Science—using search terms related to smart devices,

older adults, and chronic diseases. For smart device-related studies, we searched for smartphones, tablets, wearable devices, and fitness trackers. For chronic diseases, we searched for type 1 diabetes mellitus and type 2 (T2DM) diabetes mellitus, cardiovascular diseases, chronic lung diseases, cancer, arthritis, and the catchall term chronic disease. We searched for cardiovascular diseases including hypertension, coronary artery disease, heart failure, and stroke. Chronic lung diseases included asthma and chronic obstructive pulmonary disease (COPD). We used appropriate subject headings (eg, Medical Subject Headings) when possible (Multimedia Appendix 1 shows the search strategies). Lastly, we reviewed the bibliographies of included documents to identify additional studies.

Framework Stage 3: Selecting the Studies

To align the selected studies to the purpose of the scoping review, we used an iterative process of developing inclusion and exclusion criteria. Inclusion criteria were smart device-based mHealth studies that (1) explicitly recruited older adults, or where the average age of participants was 50 years or older, (2) aimed to support chronic disease management, (3) were published in 2010 or after, and (4) were written in English. Only articles published in 2010 or after were selected to accommodate the introduction of tablets and the wide adoption of smartphones [18]. The exclusion criteria were (1) SMS or interactive voice response-based mHealth interventions, (2) studies that validated electronic versions of scales or questionnaire forms of existing instruments, (3) smart device-based interventions for postoperative monitoring, and (4) studies that described the design, development, or usability evaluation of smartphone-based mHealth systems. A single reviewer screened for the title and abstract first, followed by a full-text review.

Framework Stage 4: Charting the Data

This stage adopted the Arksey and O'Malley framework, an iterative process of creating and revising a data extraction form to add or remove variables as the author's knowledge on the topic increased [16]. To optimize the data extraction and charting process, we created an Excel file (Excel 2011 for Mac; Microsoft Corporation) to collate the following variables that were pertinent to the aims of this research: (1) bibliographies, (2) chronic condition, (3) study design, (4) biometrics and PROMs, (5) type of self-management, (6) type of decision support, and (7) clinical and health outcomes.

Framework Stage 5: Collating, Summarizing, and Reporting the Results

The 3 substages of this stage were (1) analyzing the data, (2) reporting results, and (3) interpreting the results. For the first substage, we assessed the basic numerical analyses of the extent and distribution of the studies. The Arksey and O'Malley framework recommends adopting a theoretical framework to collate and summarize the extracted variables in a systematic manner [17]. Based on the extracted variables resulting from framework stage 4, we uncovered a high resemblance to the 3 elements of the chronic care model (CCM) [19,20] (Table 1 [19-22]). We borrowed these elements from the CCM as analytical themes to report and interpret the results. The CCM is a widely accepted framework that primarily aims to improve overall chronic care [23]. The selected 3 elements of the CCM recognize the significance of clinical information systems and technologies to better support chronic disease, thus aligning well with the purpose of this study [20]. Using a portion of the CCM allows for a structured approach to reporting results, and thereby enhances the articulation of the results to readers [16].

We could not organize results from the qualitative and mixed-methods studies within the descriptive numerical summaries due to the incommensurable nature of these two data types. Instead, we conducted thematic synthesis to capture and synthesize qualitative data of the patient experience, following the protocol outlined by Thomas and Harden [24]. Line-by-line coding, construction of descriptive themes, and development of analytical themes were completed using EPPI-Reviewer 4.0 [25].

Studies were then grouped and reported into the following categories to answer the specific research objectives from framework stage 1: (1) study characteristics, (2) the 3 elements of the CCM, (3) clinical and health impact, and (4) patient perspectives.

Table 1. Mapping and rationale for mapping to the charted variables of the chronic care model (CCM).

Extracted variables	CCM	Description	Relevance to the extracted variable
Biometrics and patient-reported outcome measures	Clinical information systems	This element in the original CCM pertains to the use of patient, care, and outcome information to gain feedback, improve practice, and develop shared care plans [20]. This idea within the context of eHealth has evolved to accommodate advanced information technologies such as electronic medical records, personal health records, smartphones, and wearable devices. [21]	This extracted variable is the information collected from smart devices. The CCM claims that the role of clinical information systems is to capture information about patients, care, and outcomes to support medical practices. Biometrics and patient-reported outcome measures are often captured and remotely sent to clinicians to support clinical practice.
Type of self-management	Self-management support	The CCM puts patients at the center of chronic disease management, and they are considered the principal caregiver. Patient education to teach necessary skills, tools to monitor symptoms, and routine assessment are integral components of self-management support [22].	Various self-management techniques that were employed by participants were extracted and categorized. These self-management techniques correspond to the acquired skills, use of tools, and assessment of accomplishments for chronic disease management.
Type of decision support	Decision support	Decision support refers to the integration of evidence-based clinical guidelines, protocols, and standards for health care providers.	This variable was extracted to identify various methods of communicating with patients to support clinical decision making.
Not applicable	Delivery system design	Delivery system design is an element that draws on the restructuring of medical practices with clearly redefined roles to support patients with chronic diseases.	No study that met the inclusion criteria directly examined the impact of altered medical practice. However, the lack of redesign of delivery systems and its impact on workload is captured in the qualitative analyses.
Not applicable	Health care organization	Health care organization is the foundational component for the model. It mainly revolves around the alignment of goals and visions across the health care organizational structure, and policies on care delivery, reimbursement, and patients.	The scope of this review study focused at the level of individual interventions rather than policy and organizational development. Thus, the nonincluded studies examined health care organization. An evaluation study may be more appropriate to assess the relationship between health care organizational change and its impact on chronic disease care for individual patients.
Not applicable	Community resources and policies	Community resources and policies refer to the need for sufficient resources within community settings (eg, exercise programs), so that primary clinics can link patients to adequate support for chronic disease management.	Smart device interventions are still at the early stages of development. The availability of community resources to support these interventions is not found within the studies.

Results

Search and Screening Results

The search of the 4 online databases in April 2016 yielded 2666 articles. We exported them to Mendeley reference management software (Mendeley for Mac version 1.17; Mendeley Ltd), where 1714 duplicates were removed. Title and abstract screening resulted in the exclusion of 821 articles, and we reviewed the remaining 131 articles in full. In total, we selected 48 articles and through review of the bibliographies identified an additional 10 potentially relevant articles. Repeating the screening cycle yielded an additional 3 articles. Figure 1 charts the detailed search, screening, and exclusion results. Ultimately, we included 51 articles (Multimedia Appendix 2 [26-76]), of which 44 were journal articles, 6 were poster abstracts, and 1 was from a conference proceedings. One poster abstract [26] and a journal article [27] described the same study.

Figure 1. Flowchart of the study selection process and the reasons for exclusion.

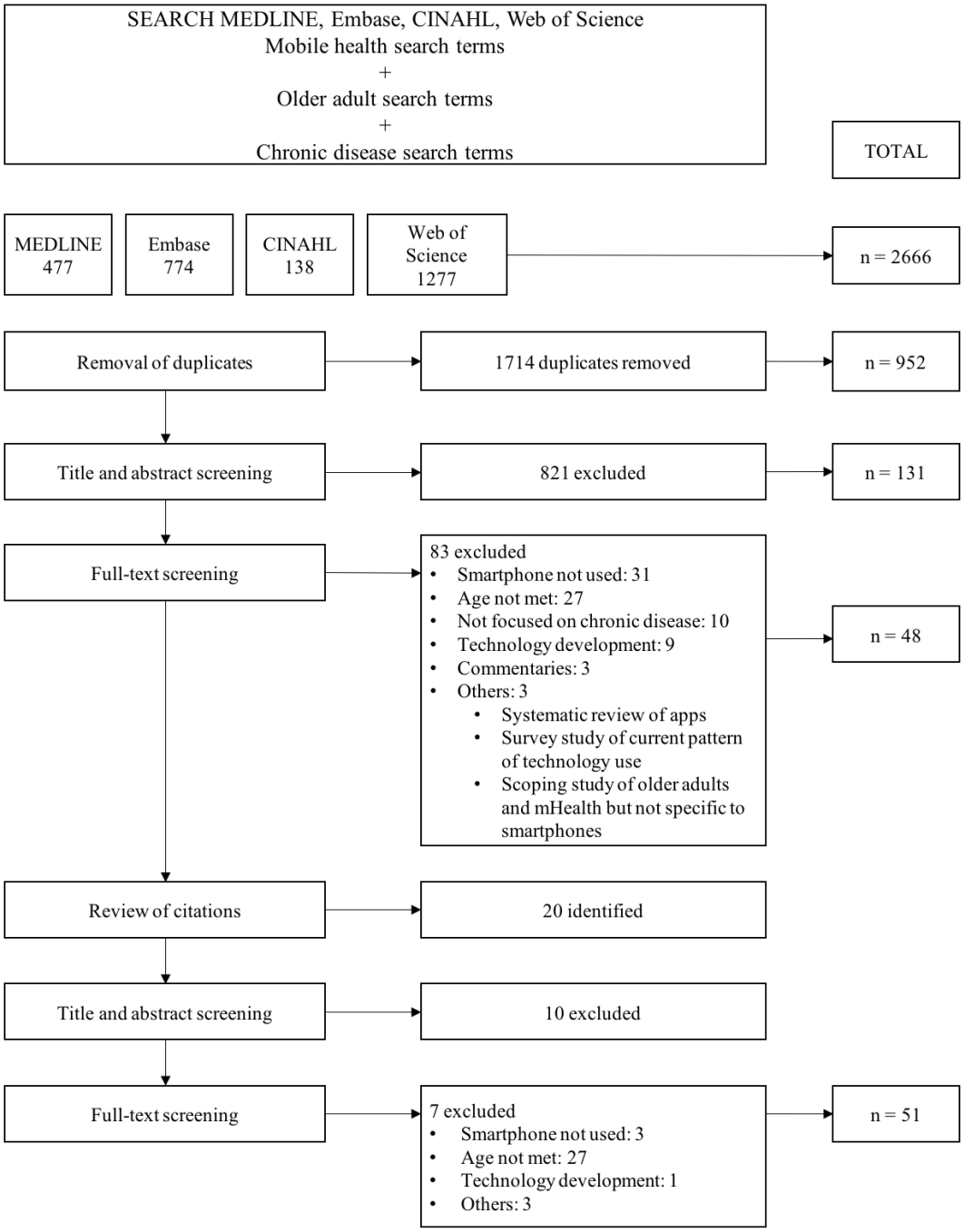
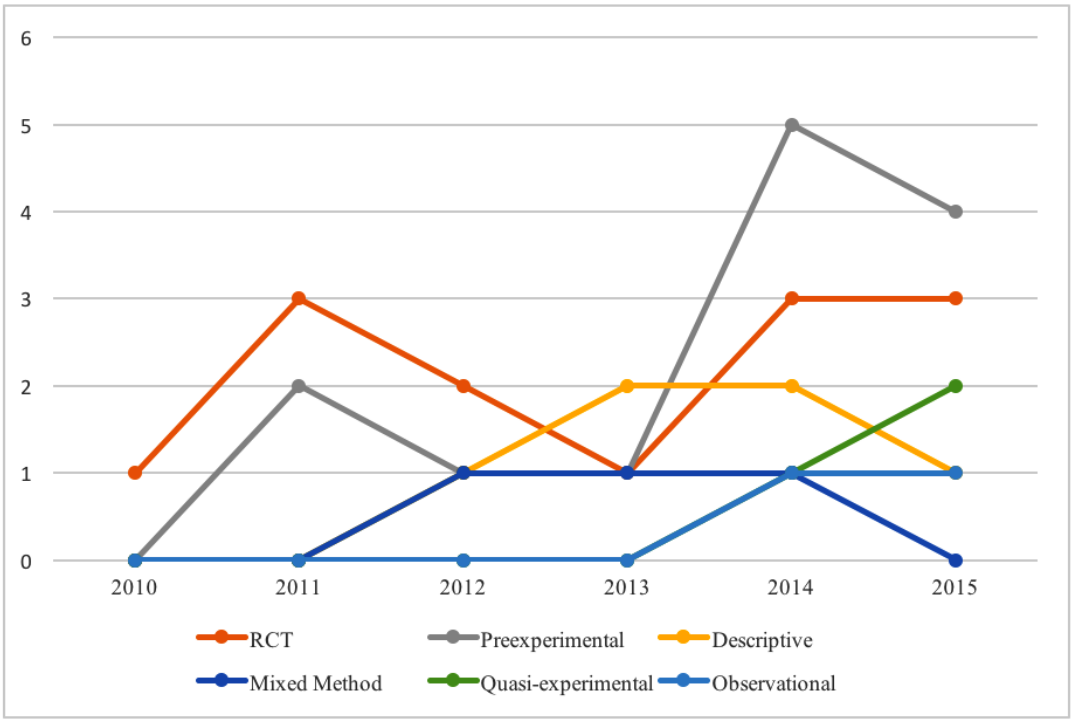


Figure 2. Type and number of study designs over time from 2010 to 2015. RCT: randomized controlled trial.



Descriptive and Numerical Summary

The sources of the studies were mainly journal articles (43/51, 84%), conference poster abstracts (6/51, 12%), and conference proceedings (2/51, 3%). The study designs varied vastly. Preexperimental design studies (16/50, 32%) that included pilot and feasibility studies employing a single-group, pretest, and posttest design were most common. We also found RCTs (13/50, 26%), qualitative studies (6/50, 12%), quasi-experimental studies (5/50, 10%), and mixed methods (4/50, 8%). Study sample size ranged from 4 to 471, and intervention duration lasted from a few hours for interviews and focus groups to 1 year. Overall, there was a pattern of increasing smart device-based research activities over time, with a noticeable spike in 2014. Figure 2 presents the yearly pattern and the different types of study designs between 2010 and 2015. The figure excludes 2016, as we included only the first 4 months, secondary data analyses (3/50, 6%), and a study protocol (1/50, 2%), as they were not primary research studies.

Of the 46 primary studies (excluding the protocol, n=1; secondary data analysis studies, n=3; and poster abstract, n=1), 16 focused on T2DM (35%), 9 focused on chronic heart failure (20%), 5 focused on COPD (11%), 4 focused on hypertension (9%), 3 focused on cancer (7%), 1 focused on asthma (2%), and 1 focused on rheumatoid arthritis (2%); 7 studies examined more than 1 chronic condition in a single study (15%) and all examined T2DM management as a major topic.

Chronic Care Model

A total of 3 extracted variables from each article were congruent with 3 of 6 elements in the CCM. The 3 extracted variables were biometrics and PROMs, type of self-management support, and type of decision support (Table 1).

Clinical Information Systems: Smart Devices Collecting Biometrics and PROMs

The smart devices were used for the collection of biometrics and PROMs—patient-measured information regarding their health status—and for the exchange of these data between health care providers and patients [77]. These tasks seamlessly align with 1 element of the CCM: clinical information systems. To explore the effective use of the smart devices as a clinical information system, this study examined the heterogeneity of the biometric data collected, and the methods of data generation and transmission.

Table 2. Number of studies collecting biometric measurements for each chronic disease.

Measurements	T2DM[a]	Heart failure	Multimorbidities	COPD[b]	Hypertension	Asthma	Total
Blood glucose	14		2				16
Body weight	4	6	2	1			13
Blood pressure	4	4	3		1		12
Step count	6		3	1			10
Heart rate		3		2	1		6
Oxygen level		1		2			3
Temperature				1			1
Peak expiratory flow rate						1	1
Electrocardiogram		1					1
Total	28	15	10	7	2	1	63

[a]T2DM: type 2 diabetes mellitus.

[b]COPD: chronic obstructive pulmonary disease.

Table 3. Number of studies collecting patient-reported outcome measures (PROMs) for each chronic disease.

PROMs	T2DM[a]	Hypertension	Cancer	Multimorbidities	Heart failure	Rheumatoid arthritis	COPD[b]	Asthma	Total
Symptoms	3	1	3		3			1	11
Medication adherence	2	2		2	1	1			8
Diet	6								6
Exercise	4			1					5
Well-being		1		1		1	1		4
Photos	2		1						3
Gait						1			1
Total	17	4	4	4	4	3	1	1	38

[a]T2DM: type 2 diabetes mellitus.

[b]COPD: chronic obstructive pulmonary disease.

A total of 40 studies used smartphones and peripheral devices to measure biometrics and PROMs [28,29,31-37,39-54,56,57,59-61,63,65,66,68-72,74,76]. We identified heterogeneity in the collection of biometrics measurements (Table 2) and PROMs (Table 3) for each chronic disease between and within chronic diseases.

Data transfer from the peripheral devices and the hosting smartphone or tablet were either manual, relying on patients to record the data on the smartphone, or automatic, via Bluetooth, near-field communication, or universal serial bus. Manual recording of the measured data created an additional burden on patients [51]. The added burden potentially influenced overall compliance with the intervention, thus altering the overall impact on chronic disease management [57]. The degree of automation of data transfer was only 62% (39/63) for biometric measurement despite technical capabilities for automation (Table 4).

Self-Management Support: Strategies for Supporting Self-Management

Another element of the CCM is self-management support. The extracted variable “types of self-management” fits well with this CCM element, as we charted the studies to understand diverse self-management support strategies that were embedded within smart device-based interventions.

We identified 8 strategies to support self-management of chronic disease through smart devices (Textbox 1) among 40 articles [27-29,31,34-37,39-54,56,57,59-61,63,65,66,68-72,74-76]. Each strategy was distinct, yet they were often highly associated with one another. For example, automated feedback sent to patients immediately after taking measurements contains information that may overlap with patient education.

Table 4. Degree of automation of biometric measurements.

Measurements	Data transfer method, n (%)		Total
	Automatic	Manual	
Blood glucose	11 (69)	5 (31)	16
Blood pressure	7 (58)	5 (42)	12
Body weight	6 (46)	7 (54)	13
Electrocardiogram	1 (100)	0 (0)	1
Oxygen saturation %	2 (67)	1 (33)	3
Peak expiratory flow rate	0 (0)	1 (100)	1
Pulse	3 (50)	3 (50)	6
Step count	9 (90)	1 (10)	10
Temperature	0 (0)	1 (100)	1
Total	39 (62)	24 (38)	63

Textbox 1. The 8 identified self-management strategies.

1. Self-monitoring: self-monitoring of the various biometrics, symptoms, medication, or healthy behaviors.
2. Patient education: education of patients pertinent to disease outcomes, self-monitoring, interpretation of measurements, benefits and risks of healthy behaviors, and medication and side effects.
3. Reminders: reminders for medication, self-monitoring, or behavior change.
4. Automated feedback: feedback content including motivational messages, educational messages, or how patients' values compare with a clinical guideline.
5. Coaching: active coaching involving structured and predefined sessions with health care providers through in-person, over-the-telephone, and virtual interactions for the purposes of education, motivation, and discussion about self-management strategies.
6. Goal setting: individualized goal setting for the treatment or behavior change.
7. Treatment plan: treatment plan outlining a protocol to follow when patients experience exacerbations of symptoms.
8. Social support: sharing of the self-management progress to engage family members and friends.

Overall, there was an inverse relationship between the number of studies and the number of self-management support strategies employed. A total of 14 studies (35%) used a single self-management support strategy and 11 studies (28%) used 3 strategies; 8 (20%), 4 (10%), and 3 (8%) articles incorporated 3, 4, and 5 types of self-management support, respectively.

The most frequently used self-management technique was self-monitoring, found in all but 2 articles [27,75] (Table 5). The second most frequently used strategy was automated feedback [37,43,48,49,51,54,56,59,63,66,68-71,76]. The nature and extent of the automated feedback varied between studies. The most frequent form of automated feedback (9/15, 60%) related to the collected measurements (eg, blood pressure) or patient progress (eg, number of steps) against the predefined parameters or goals [43,48,51,59,66,68-71]. Other feedback included motivational, self-care, and educational messages. Patient education was another frequently used type of self-management support, and the extent of patient education differed considerably. Of the 13 smart device systems, 9 had a dedicated user interface for patient education; 7 of these included videos and audio files on topics ranging from instructions on proper use of peripheral devices, to general information about the disease, to proper exercise techniques [27,36,39,43,51,65,75]. Two studies merely provided a brief educational session before the start of the research study and supplemented the session with pamphlets and information sheets [47,66]. One study that aimed to enhance medication adherence provided information about each medication the patient was taking [52].

Table 5. Frequency of self-management support strategies employed within 40 studies.

Types of self-management support	Number of studies
Self-monitoring	38
Automated feedback	15
Patient education	13
Reminder	8
Coaching	8
Goal setting	6
Social support	2
Treatment plan	1

Decision Support

The CCM describes decision support as a function of the advanced clinical information system. The purpose of a decision support system (DSS) is to assist health care providers to adhere to an evidence-based clinical guideline. Within the context of smart device-based interventions, the support was provided directly to patients rather than through health care providers. Patients received support for decision making about self-management from diverse sources, categorized based on the medium of delivery as (1) virtual (a form of online communication, usually messaging), (2) in person, (3) by telephone, or (4) through a DSS.

A total of 23 articles documented aid in the decision-making process for patients. Of these, 14 studies (61%) integrated 1 type of decision support mechanism [28,34,37,39,42,47,54,63,65,66,68,70,72,76] and 8 studies (35%) integrated more than 1 type of decision support mechanism [36,41,43,45,48,56,59,69]. These interventions were designed to delay the involvement of health care providers to reduce the burden of monitoring patients. In total, 12 studies (52%) integrated a DSS [37,41,43,47-49,54,56,59,63,69,70]. The comprehensiveness and setup of the DSSs varied considerably: 3 smart device system-implemented DSSs not only determined the severity of the disease, but also recommended adjusting medications or starting a medication regimen based on the patient's biometrics and PROMs [43,63,70]. Of 4 DSSs operating on evidence-based guidelines [37,43,47,70], 3 recommended medication adjustments. Another 4 studies allowed health care providers to determine the individualized parameters for each patient, and the DSS generated alerts based on these predefined values [48,49,59,69].

Clinical and Health Impact

We assessed the impact of smart device-based interventions on patients' clinical and health outcomes. Since this was not a systematic review, we did not assess the quality of the evidence; rather the review was aggregated and observational.

Of the 51 studies, 10 RCTs and 2 quasi-experimental studies examined clinical outcomes [36,42,45,47,48,50,56,59,63,69,71,76]. The degree of clinical impact varied for different conditions and each clinical measure (Table 6).

The most frequently reported health improvement was HbA_{1c} control among patients with diabetes, while no superior clinical impact on blood lipid profile was observed [36,42,45,48,56,71]. Patients with asthma showed improvements in the percentage predicted peak expiratory flow rate and forced expiratory volume in the first second of expiration, but these findings were based on 1 RCT [47]. Patients with gastric cancer who had undergone gastrectomy and used smart devices to manage symptoms and nutritional intake were able to retain their body weight significantly better than the control group [76]. No clinical improvement was reported for heart failure and COPD intervention studies [59,62,69].

Impact on Quality of Life

Many smart device-based chronic disease management programs aimed to manage symptoms and improve quality of life rather than solely focusing on improving clinical outcomes. Quality of life was assessed in 9 studies, and 6 assessment instruments were used [28,37,42,47,49,54,59,66,68]. Only the 36-Item Short Form Health Survey (SF-36) and the EuroQoL Five Dimensions Questionnaire (EQ-5D) were each used in 2 studies (Table 7).

A total of 3 RCTs reported significant improvements in the quality of life for patients with heart failure and asthma [37,47,59]. An RCT for patients with diabetes and COPD documented a mixed result of improvements in the physical component, but not in the mental component [66]. Another diabetes management intervention also reported a mixed result on quality of life. Patients reported less physical pain by the end of 4 weeks but no improvements in other domains of the SF-36 [54].

Table 6. Number of clinical outcomes of randomized controlled trials and quasi-experimental design studies of smart device-based mHealth interventions.

Chronic diseases (measurements)	Significant outcomes	Nonsignificant outcomes
T2DM[a] (hemoglobin A_{1c})	3	1
T2DM (blood pressure)	1	4
T2DM (lipids)	0	5
HF[b] (brain natriuretic peptide, LVEF[c])	0	2
HF (self-care activities)	1	1
HF (lipids, blood pressure, weight, waist circumference)	0	1
Chronic obstructive pulmonary disease (dyspnea, fatigue level)	0	1
Asthma (peak expiratory flow rate, FEV_1 % pred[d])	1	0
Cancer (body weight)	1	0
Hypertension (blood pressures)	1	0
Total	8	15

[a]T2DM: type 2 diabetes mellitus.
[b]HF: heart failure.
[c]LVEF: left ventricular ejection fraction.
[d]FEV_1 % pred: percentage predicted forced expiratory volume in the first second of expiration.

Table 7. Smart device-based mHealth interventions and impact on quality of life.

Study (first author, date, reference no.)	Study design	Morbidity	Impact on quality of life	Measurement tool
Anglada-Martinez, 2016 [28]	Preexperimental	Multimorbidities	Not significant	EQ-5D[a]
Hägglund, 2015 [37]	RCT[b]	Heart failure	Significant	KCCQ[c]
Karhula, 2015 [42]	RCT	T2DM[d] and heart failure	Not significant	SF-36[e]
Liu, 2011 [47]	RCT	Asthma	Significant	SF-12[f]
Maguire, 2015 [49]	Preexperimental	Cancer	Not significant	FACT-L[g]
Quinn, 2015 [54]	Preexperimental	T2DM	Mixed	SF-36
Seto, 2012 [59]	RCT	Heart failure	Significant	MLHFQ[h]
van der Weegen, 2015 [66]	RCT	T2DM and COPD[i]	Mixed	RAND-36[j]
Verwey, 2014 [68]	Mixed methods	COPD	Significant	EQ-5D

[a]EQ-5D: EuroQoL Five Dimensions Questionnaire.
[b]RCT: randomized controlled trial.
[c]KCCQ: Kansas City Cardiomyopathy Questionnaire.
[d]T2DM: type 2 diabetes mellitus.
[e]SF-36: 36-Item Short Form Health Survey.
[f]SF-12: 12-Item Short Form Health Survey.
[g]FACT-L: Functional Assessment of Cancer Therapy – Lung.
[h]MLHFQ: Minnesota Living with Heart Failure Questionnaire.
[i]COPD: chronic obstructive pulmonary disease.
[j]RAND-36: RAND 36-Item Health Survey.

Table 8. Descriptive and analytical themes identified in 5 smart device-based mHealth research studies.

Themes	Hallberg, 2016 [38]	Maguire, 2015 [49]	Nes, 2012 [51]	Verwey, 2014 [68]	Williams, 2014 [73]
Inappropriate target patients					
No symptoms and stable condition	Yes	N/A[a]	N/A	Yes	N/A
Insufficient training					
Insufficient knowledge on how to use the system	Yes	Yes	N/A	Yes	N/A
Insufficient knowledge on how to interpret data	Yes	N/A	Yes	N/A	Yes
Anxiety toward technology use	N/A	N/A	N/A	N/A	Yes
Recognition of benefits					
Increased self-awareness of symptoms and disease conditions	Yes	Yes	Yes	Yes	Yes
Increased motivation to upkeep chronic disease management	Yes	Yes	Yes	Yes	Yes
Increased knowledge about chronic disease and management	Yes	N/A	N/A	N/A	Yes
Increased involvement and engagement in chronic disease management	Yes	N/A	Yes	N/A	Yes
Improved chronic disease management behaviors and outcomes	N/A	Yes	Yes	N/A	Yes
Utility as a communication tool	Yes	Yes	N/A	N/A	Yes
A sense of connectedness and reassurance					
Feeling assured	N/A	Yes	N/A	N/A	Yes
Reduced uncertainty around chronic disease management	N/A	Yes	N/A	N/A	Yes
Personalized feedback, advice, and messages	N/A	Yes	Yes	Yes	N/A
System issues diminish motivation					
Usability issues	Yes	N/A	Yes	Yes	N/A
Perceived ease of use	Yes	Yes	N/A	N/A	Yes
Feeling burdened	N/A	N/A	Yes	N/A	Yes

[a]N/A: not applicable.

Thematic Synthesis

For the thematic analysis, we chose 5 smart device-based mHealth research studies that aimed to explore patient perspectives [38,49,51,68,73]. Initially, we constructed 16 descriptive themes from line-by-line coding. We condensed the descriptive themes into 5 analytical themes (Table 8), then further categorized them as the facilitators and barriers to successful adoption by older adults.

Facilitators

Theme 1: A Sense of Connectedness and Reassurance

Patients understood that their measurements and recordings were observed by health care providers. A sense of reassurance was offered when patients received feedback from clinicians.

> *...the rapid feedback by health professionals in response to reported symptoms...it just keeps your morale up...(pg E43).* [49]

A sense of connectedness and feeling of reassurance was offered regardless of whether health care providers were monitoring the system or not.

> *Yet the virtual link offered by mHealth intervention appeared to reassure patients and gave a sense of continuity of care...the sharing of patients' self-monitoring data with the research nurse, even though this was infrequent and did not replace current care (pg e395).* [73]

Having the option to contact health care providers at any time reduced uncertainties that arose when self-managing symptoms and evoked a sense of reassurance.

> *...reduced uncertainty experienced by the patients, particularly at times when they were at home and were unsure as to whether they should contact health professionals or not (pg E43).* [49]

Theme 2: Recognition of the Benefits of Using Smart Devices

Many participants from multiple studies commented on the increased awareness, motivation, and engagement in chronic disease management owing to the use of smart device systems. Patients frequently reflected on their self-management behaviors while filling out diaries for PROMs, noting it enhanced their sensitivity to and awareness of their symptoms.

> *...filling in the diaries gave them better insight into their diabetes, increased their coping and self-management strategies...increasing their honesty in answering the diary questions (pg 389-90).* [51]

> *...reviewing their condition and how they felt not just over a number of days but also within a 24-hour period...As patients were answering questions about their symptoms as well as monitoring their oxygen saturation on daily basis, they felt encouraged to think more about how they were feeling each day and throughout the day (pg e395).* [73]

Patients were more motivated to comply with mHealth interventions by achieving predefined goals. Also, patients recognized the significance of complying with an mHealth program when they experienced the relationship between their self-management behaviors and its impact on their health.

> *They felt encouraged to be more active and mentioned three aspects [that were positive about the intervention]: the awareness of their physical activity performance, the stimulating effect of the daily target goal and the positive effect on self-efficacy (pg 31).* [68]

> *This became evident to those who missed their antihypertensive medication at some point and then personally detected that their blood pressure had gone up the same day (pg 144).* [38]

Barriers

Theme 3: Incompetent Recruitment Strategy

Smart device-based systems did not offer much value for some recruited participants, although this varied given patients' chronic disease management status. These patients consistently commented on the lack of value the intervention added to their management strategies. In one study that targeted hypertensive patients, the authors described the experience of those who were noncompliant with the mHealth program:

> *...patients who did not perceive any symptoms or who had stable blood pressure found some questions to be less relevant (pg 143).* [38]

Another study that promoted physical activity among patients with COPD identified the reasons for lower ratings when asked if the intervention had a positive effect on their physical activity level.

> *A total of 12 patients were positive about the effect of the intervention on their physical activity performance and five patients were neutral about it; the latter were patients who were already sufficiently active (pg 31).* [68]

Smart device systems that visually represented patients' longitudinal progress were viewed as less useful, since patients with stable conditions were well aware without seeing the graph.

> *...they did not consider it necessary to look at the graphs, as their answers and values were quite similar without significant variation (pg 143).* [38]

Although many studies had explicit eligibility criteria to recruit appropriate patients, they did not filter out patients who were already sufficiently managing chronic diseases. Alternatively, the targeting of unstable patients for smartphone-based mHealth interventions could pose a threat to their safety.

Theme 4: Insufficient Training

The theme of insufficient training encompasses 2 distinct issues. First, it refers to a lack of training and follow-up assistance for using smart device systems.

> *Some of the participants commented on how they were never trained on using this component of the system (pg E43).* [49]

Moreover, some smart device systems were set up to operate with computers, expanding the need for technical training. Participants did not pursue extra help and often did not use a portion of, or an entire, system. This was evident in a study where patients had an option to view their progress on a computer.

> *...simply logging into the computer to look at the graph was a problem in itself because they had forgotten how to do this. Therefore, some patients did not in fact look at the graphs themselves in their own home...(pg 143).* [38]

Another study highlighted the need for additional training, noting that not all participants learned to use the system at the same pace.

> *A total of six patients needed some extra advice about how to log in, which was given to them (pg 32).* [68]

Second, there was a lack of patient education to support the self-management and interpretation of measured biometrics.

> *The participants had doubts especially about how to interpret these questions [referring to questions for daily diaries (pg 390).* [51]

> *...patients were uncertain how to interpret oxygen levels (pg e396).* [73]

The lack of knowledge on how to interpret questions and the measured biometrics and PROMs led to inaccurate reporting of their symptoms or left participants uncertain about their conditions and progress. Insufficient and inadequate levels of training had a leaching effect on nurses and therapists, with an unanticipated increase in workload due to issues such as longer consultations [51,68].

Theme 5: Diminished Motivation Due to Usability and Technical Issues

Usability and technical issues were prevalent across the studies, and they led to frustration among older adults and had a detrimental effect on patient motivation to continue with the study.

> *Although most of the patients were positive about the tool, the motivation of some patients dropped when technical problems occurred (pg 32).* [68]
>
> *Participants reported technical problems as frustrating and demotivating (pg 391).* [51]

Most studies reported varying degrees of technical and usability issues with smart systems. These issues are unavoidable in any technology-driven intervention, but no study described a risk mitigation strategy in case of technology failure. Patients, more and more, play a critical role in managing lifelong illnesses, and smart device-based systems can play a central role in delivering necessary support. As a result, these issues posed a significant threat to effective use of the tool, thus hindering chronic disease management for some older adults.

Discussion

Limitations of Smart Device-Based Research

This scoping review explored the extent of the use of smartphones, tablets, and peripheral devices, the diverse self-management support strategies, and the nature of decision support and DSSs for managing chronic diseases among older adults. Further, this study illustrated older patients' experience with using such technologies for their care.

Lack of Gerontological mHealth Research

Overall, smart device-based research for older adults began to rise with a noticeable increase in 2014. This is similar to previous literature findings [11]. Despite the increased research activities, only 3 studies reviewed here explicitly targeted older adults, revealing a lack of gerontological smart device-based studies [50,54,74]. This is a lost opportunity as more and more older adults are displaying interest in using smartphones and tablets for obtaining health information [78]. This trend is expected to grow as the proportion of older adults living with chronic conditions is increasing.

Dominance of Preexperimental Study Design

mHealth studies have long been criticized for the dominance of preexperimental studies that labeled themselves as pilot, feasibility, and field-test studies [5,12], and this trend continued among smart device-based mHealth research. These types of studies lacked a control group and are subject to threats to internal validity [79]. Consequently, these studies offer limited knowledge about the effectiveness of smart device-based interventions in producing meaningful clinical outcomes compared with RCTs and quasi-experimental studies [5]. However, a strong dominance of pilot studies and an increasing number of RCTs is indicative of overall progress in this topic.

Narrow mHealth Research Scope

Diabetes and hypertension management has been the dominant subject for traditional mHealth research, whereas other chronic disease management received very limited attention [2,4]. This pattern strongly continued among smart device-based mHealth research. More chronic diseases have been explored using smart devices, including heart failure, cancer, asthma, and multichronic disease management. This expansion of the scope of mHealth research is a welcome change, and these studies have provided deeper understanding of the potential for smart devices in chronic disease management for older adults. However, the scope of mHealth research should expand further as technological capabilities of smart devices increase, but it was too early for this scoping review to detect meaningful diversification.

Inadequate Support for Participants

The literature provides very limited information about the nature of training (ie, number of participants per training session; number of hours) and the content of the training (ie, how to use a smartphone; how to interpret blood glucose levels). However, the thematic synthesis of qualitative studies unveiled an issue with research practice in regard to training of participants. Moreover, there was no consistent structure for ongoing support for participants. No standard practice for participant training on technology-driven interventions had many participants discouraged from using smart devices to their full potential and often led to frustration among health care providers.

Insufficient Evidence for Health and Clinical Outcomes

The evidence for smart device-enhanced clinical and health outcomes was insufficient and mostly inconclusive. Smart device-based interventions for older patients with diabetes demonstrated its effectiveness in reducing the HbA_{1c} level, but it did not demonstrate superior clinical outcomes than the traditional mHealth interventions [80]. Similarly, minimal evidence was documented for improvement in chronic cardiovascular and chronic respiratory disease management. Many studies only reported recognizable clinical benefits among a particular group of patients through the use of subanalyses and emphasized the potential these new technologies hold for the future.

Implications for Future Research

Need for Evidence-Based Design of Smart Device-Based Programs

We observed disjointed chronic disease care processes across studies, and even within the studies that focused on the same chronic conditions. Additionally, the lack of research activities that focused on the geriatric population is of great concern. The collection of a wide range of patient biometrics and PROMs, and the use of many different self-management strategies indicate the absence of standards for the evidence-based design of smart device-based programs for older adults.

Heterogeneity in intervention design is not unique to smart device-based interventions. Heterogeneity in the content of interventions for different chronic conditions, and even within a single chronic condition, was also reported in a scoping review that examined nontechnology-derived chronic disease self-management programs [81]. This study indicated that no single design would suffice for the design of smart device-based programs but, instead, the design of interventions should ensure that components of such programs are evidence based.

Research activities for older adults had a skewed focus on validating the effectiveness of a handful of components such as self-monitoring and automated feedback, which had been extensively researched through traditional mHealth interventions

through the use of text messaging and self-monitoring devices [2,4,5]. These studies demonstrated limited clinical effectiveness for older adults for managing chronic diseases [2,4,5]. This finding is in line with behavioral change interventions for older adults in weight management and physical activity delivered via text messages and self-monitoring [2,82]. In contrast, mHealth intervention studies for younger populations in smoking cessation through text messaging demonstrated a more than doubled smoking cessation rate [2,4]. More importantly, diabetes management interventions via text messaging for adolescents and young adults demonstrated increased blood glucose monitoring frequency and clinically significant improvement in HbA_{1c} level [4]. Such a pattern indicates that smart device-based interventions for older adults need to exploit the technical advantages for testing and expanding new and novel strategies for self-management. Example strategies that used the new features of smart devices are social support through the use of social network services to involve others [83] and goal setting for successful behavior changes through visually presenting progress [83,84]. These are only a few examples of evidence-based strategies uncovered in this scoping review; future smart device-based interventions should explore other evidence-based self-management support strategies and validate their generalizability, thus leading to the development of more evidence-based chronic disease management programs.

Patient Decision Support Systems and Data-Driven Health Care

The role of DSSs within the context of smart device-based interventions is to support patients in their clinical decision-making needs, thereby increasing patient autonomy. Smart device-based interventions often delegate the responsibilities of clinical practices to patients with support from patient DSSs, but the degree of delegation varied widely. One study employed a DSS that initiated a medication regimen for patients based on their reported biometrics and PROMs [47]. Some studies developed simple rule-based DSSs that determined the health status of a patient automatically, and clinicians were only involved when a patient needed closer examination [43,48,49,56,59]. On the other hand, some DSSs were less intrusive by providing motivational messages, educational contents, and self-care advice [63,69].

Use of DSSs within smart device-based interventions raises multiple issues that warrant further research. First, the complete delegation of the clinical decision-making process to patients—albeit with validated DSSs—will need extensive research on its safety, efficacy, and implications for chronic care management and patient well-being. This may be an appropriate approach for simple situations where there is a single decision option with suitable patient education. However, many medical decisions will benefit from clinician involvement to allow careful examination of available decision options that can lead to informed shared decision making [85]. Some studies attempted to achieve this by setting up rules around when a clinician should be involved or by setting up regular scheduled supervision by clinicians. Too little interaction with clinicians or too much delegation of clinical responsibilities created uncertainty and anxiety among patients. Qualitative analyses demonstrated that online supervision was deemed sufficient to provoke the sense of connectedness (ie, online communication) to reassure patient confidence. The use of patient DSSs along with scheduled clinician supervision needs further research to optimize its impact on patient safety, patient autonomy, and resource utilization (ie, clinician time). Second, minimally intrusive DSSs that provide motivational messages and educational contents need further verification for their effectiveness in supporting self-management. Some research studies have investigated this topic, but the conclusions reported are insufficient, and the need for tailored messages was noted [38,62].

Expanding Practice Through the Use of Innovative Technologies

Overall, the adoption of innovative technologies was not well realized within smart device-based research. One aspect that was examined in detail was the diversity and the methods of gathering biometrics and PROMs. In particular, the automation of data generation and transmission was of significant importance for improving chronic disease care because it influenced older adults' compliance with smart device-based interventions [34,42]. For example, the impact of automation was well documented in one study in which patients had to wear an accelerometer (automated) and fill in diaries every day (manual): the mHealth compliance rate was much higher with accelerometer use [63]. The effectiveness of automation in improving mHealth compliance was supported by other studies, especially for older adults [34,36]. However, over one-third of the biometrics were obtained through analog medical equipment that did not automatically transfer the data, and compliance was noted to be lower. Use of peripheral devices that can automate data generation and transmission is recommended for future smartphone-based research for older adults.

While the automation of data collection will benefit smart device-based mHealth research, we should look beyond this for more innovative solutions. The topic of wearable technologies for health care is emerging and receiving growing attention [86,87]. A recent industry report that scanned the consumer-grade wearable device adoption level indicated almost as high a level of adoption among older US adults as among their younger counterparts [88]. Wearable technologies have not been explored for chronic disease management for older adults. Continuous monitoring of many vital signs, along with the simplicity of devices, provides great potential for chronic disease management, especially for older adults [89]. Early research protocols and findings from studies that used smart wearable devices to support chronic disease management among older adults have begun to emerge, indicating that mHealth and ubiquitous health research practices are expanding [90,91].

Use of multiple sensors and continuous monitoring generate an enormous amount of data, which has opened up opportunities for advanced data analytics. Data analytics has demonstrated its effectiveness in medicine, from diagnostics, clinical decision support, and cardiac event detection [92-96]. Use of advanced data analytics has been integrated into mHealth studies, albeit very limited, for general health promotion and prevention research. One study used smartphones to generated personalized,

contextualized, and actionable health advice and recommendations based on activity and nutrition intake [97]. Another study employed a neural network to predict falls based on kinematic parameters collected from a smart wearable device [98]. Future mHealth studies should explore the utility of data analytic techniques, potentially using large amount of data from continuous monitoring from smart wearable devices for detecting exacerbation among patients with COPD [99], irregular heart beat in cardiovascular disease [92], and dangerous spikes of blood glucose levels [100,101], to name a few, without resorting to clinicians. Ultimately, new innovative technologies should improve chronic disease care, increase the autonomy of patients, prolong the independence of older adults, and ensure overall well-being rather than being burdensome on patients' daily lives.

Study Limitations

This scoping study was subject to publication bias. Despite our effort to be as inclusive as possible by searching 4 separate reference databases, studies with negative findings may have been underrepresented. Another limitation of the study was the restrictive search of online reference databases and exclusion of gray literature. We excluded gray literature from this scoping review to balance the feasibility of this study with the available resources. This study was conducted by a single reviewer, which may have caused biases in selecting and screening of search results. Given the nature of the scoping review, this study did not synthesize evidence to determine the effectiveness of smart device-based tools. Instead, it captured the diversity of the available literature with its varied objectives, interventions, populations, and settings. Consequently, this study was more exploratory and suggestive of future research directions.

Conclusions

Effective care of older adults with chronic conditions is a growing global priority. The advances in smartphones, wearables, and other smart devices align well with the developing interests of older adults to integrate technologies into their health care. Future smart device-based mHealth interventions should focus on implementing evidence-based strategies into the research design, exploring more powerful and reliable data-driven DSSs, and using innovative technologies to enhance and expand mHealth practice. To unlock the potential of rapidly advancing technologies for the aging population, future mHealth research should embrace innovation and look to develop evidence-based chronic care programs for older adults.

Acknowledgments

This work was supported by a grant from the AGE-WELL Network of Centres of Excellence.

Authors' Contributions

BK and JL formulated and designed the study topic. BK participated in designing, coordinating, and conducting the study and drafting the manuscript. JL contributed to the development and refinement of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Final search queries for online reference databases.

[PDF File (Adobe PDF File), 34KB - mhealth_v5i5e69_app1.pdf]

Multimedia Appendix 2

Description of the 51 included articles.

[PDF File (Adobe PDF File), 48KB - mhealth_v5i5e69_app2.pdf]

References

1. Akter S, Ray P. mHealth - an ultimate platform to serve the unserved. Yearb Med Inform 2010:94-100. [Medline: 20938579]
2. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
3. Kay M, Jonathan S, Marina T. mHealth: new horizons for health through mobile technologies. Geneva, Switzerland: World Health Organization; 2011 Jun. URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2011-12-08] [WebCite Cache ID 63mBxLED9]
4. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]

5. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst Rev 2012;12:CD007459. [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]
6. Carrasco MP, Salvador CH, Sagredo PG, Márquez-Montes J, González de Mingo MA, Fragua JA, et al. Impact of patient-general practitioner short-messages-based interaction on the control of hypertension in a follow-up service for low-to-medium risk hypertensive patients: a randomized controlled trial. IEEE Trans Inf Technol Biomed 2008 Nov;12(6):780-791. [doi: 10.1109/TITB.2008.926429] [Medline: 19000959]
7. Kiselev AR, Gridnev VI, Shvartz VA, Posnenkova OM, Dovgalevsky PY. Active ambulatory care management supported by short message services and mobile phone technology in patients with arterial hypertension. J Am Soc Hypertens 2012;6(5):346-355. [doi: 10.1016/j.jash.2012.08.001] [Medline: 22995803]
8. Bobrow K, Farmer AJ, Springer D, Shanyinde M, Yu L, Brennan T, et al. Mobile phone text messages to support treatment adherence in adults with high blood pressure. Circulation 2016 Feb 09;133(6):592-600 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.017530] [Medline: 26769742]
9. Mosa ASM, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012;12:67 [FREE Full text] [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]
10. 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;14(5):e128 [FREE Full text] [doi: 10.2196/jmir.1994] [Medline: 23017375]
11. 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;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
12. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
13. Rainie L, Perrin A. Technology adoption by baby boomers (and everybody else). Washington, DC: Pew Research Center; 2016 Mar 22. URL: http://www.pewinternet.org/2016/03/22/technology-adoption-by-baby-boomers-and-everybody-else/ [accessed 2016-12-07] [WebCite Cache ID 6maHvfXA2]
14. Smith A. U.S. smartphone use in 2015. Chapter one: a portrait of smartphone ownership. Washington, DC: Pew Research Center; 2015 Apr 01. URL: http://www.pewinternet.org/2015/04/01/chapter-one-a-portrait-of-smartphone-ownership/ [accessed 2015-07-08] [WebCite Cache ID 6ZsC8HUFA]
15. Wang J, Wang Y, Wei C, Yao NA, 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]
16. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
17. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
18. Device ownership over time. Washington, DC: Pew Research Center; 2014 Jan. URL: http://www.pewinternet.org/data-trend/mobile/device-ownership/ [accessed 2014-09-23] [WebCite Cache ID 6SoSK3XJP]
19. Bauer A, Rue T, Keppel G, Cole A, Baldwin L, Katon W. Primary care patients' use of mobile health (mHealth) tools. J Gen Intern Med 2014 Apr;29(Suppl 1):180-181. [doi: 10.1007/s11606-014-2834-9]
20. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. Milbank Q 1996;74(4):511-544. [Medline: 8941260]
21. Gee PM, Greenwood DA, Paterniti DA, Ward D, Miller LMS. The eHealth enhanced chronic care model: a theory derivation approach. J Med Internet Res 2015;17(4):e86 [FREE Full text] [doi: 10.2196/jmir.4067] [Medline: 25842005]
22. DiBello KK, Boyar KL, Abrenica SC, Worral PS. The effectiveness of text messaging programs on adherence to treatment regimens among adults aged 18 to 45 years diagnosed with asthma: a systematic review protocol. JBI Database Syst Rev Implement Rep 2013;11(8):170-185. [doi: 10.11124/jbisrir-2013-950]
23. Grover A, Joshi A. An overview of chronic disease models: a systematic literature review. Glob J Health Sci 2014 Oct 29;7(2):210-227 [FREE Full text] [doi: 10.5539/gjhs.v7n2p210] [Medline: 25716407]
24. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008;8:45 [FREE Full text] [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
25. EPPI Centre. EPPI-Reviewer 4. London, UK: Evidence for Policy and Practice Information and Co-ordinating Centre, Social Science Research Unit, Institute of Education, University of London; 2016. URL: http://eppi.ioe.ac.uk/cms/Default.aspx?alias=eppi.ioe.ac.uk/cms/er4 [accessed 2017-05-15] [WebCite Cache ID 6qU8kJW7V]
26. Clark R, Fredericks B, Adams M, Atherton J, Howie-Esquivel J, Dracup K, et al. Development of a culturally specific heart failure self-care iPad teaching tool for Aboriginal and Torres Strait Islanders. Heart, Lung Circ 2013 Jan;22(suppl 1):S251. [doi: 10.1016/j.hlc.2013.05.599]

27. Clark RA, Fredericks B, Buitendyk NJ, Adams MJ, Howie-Esquivel J, Dracup KA, et al. Development and feasibility testing of an education program to improve knowledge and self-care among Aboriginal and Torres Strait Islander patients with heart failure. Rural Remote Health 2015;15(3):3231 [FREE Full text] [Medline: 26390941]
28. Anglada-Martínez H, Martin-Conde M, Rovira-Illamola M, Sotoca-Momblona JM, Sequeira E, Aragunde V, et al. Feasibility and preliminary outcomes of a web and smartphone-based medication self-management platform for chronically ill patients. J Med Syst 2016 Apr;40(4):99. [doi: 10.1007/s10916-016-0456-y] [Medline: 26872781]
29. Arsand E, Tatara N, Østengen G, Hartvigsen G. Mobile phone-based self-management tools for type 2 diabetes: the few touch application. J Diabetes Sci Technol 2010 Mar;4(2):328-336 [FREE Full text] [Medline: 20307393]
30. Bengtsson U, Kasperowski D, Ring L, Kjellgren K. Developing an interactive mobile phone self-report system for self-management of hypertension. Part 1: patient and professional perspectives. Blood Press 2014 Oct;23(5):288-295 [FREE Full text] [doi: 10.3109/08037051.2014.883203] [Medline: 24564289]
31. 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]
32. Cohen J, Peeples M, Iyer A, Shomali M. Mobile prescription therapy: the potential for patient engagement to enhance outcomes. 2015 Presented at: 75th Scientific Session of the American Diabetes Association; Jun 4-9, 2015; Boston, MA, USA p. A230 URL: https://www.welldoc.com/images/uploads/ADA_poster_2015_Final.pdf
33. Dang S, Oropesa L, Byrne M, Gutt M, Andrade F, Guanipa C. Race/ethnic disparities in weight and glycemia in older adults receiving lifestyle interventions via peer-leaders with or without mobile enhancement for diabetes prevention and management. J Am Geriatr Soc 2014;62(Suppl 1):C140.
34. Ding H, Karunanithi M, Kanagasingam Y, Vignarajan J, Moodley Y. A pilot study of a mobile-phone-based home monitoring system to assist in remote interventions in cases of acute exacerbation of COPD. J Telemed Telecare 2014 Apr;20(3):128-134. [doi: 10.1177/1357633X14527715] [Medline: 24643954]
35. Goldstein CM, Gathright EC, Dolansky MA, Gunstad J, Sterns A, Redle JD, et al. Randomized controlled feasibility trial of two telemedicine medication reminder systems for older adults with heart failure. J Telemed Telecare 2014 Sep;20(6):293-299. [doi: 10.1177/1357633X14541039] [Medline: 24958355]
36. 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;17(7):e178 [FREE Full text] [doi: 10.2196/jmir.4112] [Medline: 26199142]
37. Hägglund E, Lyngå P, Frie F, Ullman B, Persson H, Melin M, et al. Patient-centred home-based management of heart failure. Findings from a randomised clinical trial evaluating a tablet computer for self-care, quality of life and effects on knowledge. Scand Cardiovasc J 2015 Aug;49(4):193-199. [doi: 10.3109/14017431.2015.1035319] [Medline: 25968968]
38. Hallberg I, Ranerup A, Kjellgren K. Supporting the self-management of hypertension: patients' experiences of using a mobile phone-based system. J Hum Hypertens 2016 Feb;30(2):141-146 [FREE Full text] [doi: 10.1038/jhh.2015.37] [Medline: 25903164]
39. Hardinge M, Rutter H, Velardo C, Shah SA, Williams V, Tarassenko L, et al. Using a mobile health application to support self-management in chronic obstructive pulmonary disease: a six-month cohort study. BMC Med Inform Decis Mak 2015 Jun 18;15:46 [FREE Full text] [doi: 10.1186/s12911-015-0171-5] [Medline: 26084626]
40. Hunt CW, Sanderson BK, Ellison KJ. Support for diabetes using technology: a pilot study to improve self-management. Medsurg Nurs 2014;23(4):231-237. [Medline: 25318336]
41. Johnston NW, Lambert K, Hussack P, Gerhardsson de Verdier M, Higenbottam T, Lewis J, et al. Detection of COPD exacerbations and compliance with patient-reported daily symptom diaries using a smart phone-based information system [corrected]. Chest 2013 Aug;144(2):507-514. [doi: 10.1378/chest.12-2308] [Medline: 23519329]
42. Karhula T, Vuorinen A, Rääpysjärvi K, Pakanen M, Itkonen P, Tepponen M, et al. Telemonitoring and mobile phone-based health coaching among Finnish diabetic and heart disease patients: randomized controlled trial. J Med Internet Res 2015;17(6):e153 [FREE Full text] [doi: 10.2196/jmir.4059] [Medline: 26084979]
43. Kim EK, Kwak SH, Baek S, Lee SL, Jang HC, Park KS, et al. Feasibility of a patient-centered, smartphone-based, diabetes care system: a pilot study. Diabetes Metab J 2016 Jun;40(3):192-201 [FREE Full text] [doi: 10.4093/dmj.2016.40.3.192] [Medline: 27098508]
44. Kim EK, Kwak SH, Park KS, Jang HC, Cho YM. OP4 Efficacy of mobile diabetes care based on a newly developed patient decision support system (PDSS). Diabetes Res Clin Pract 2014 Nov;106(January):S2-S3. [doi: 10.1016/S0168-8227(14)70210-4]
45. Kim H, Choi W, Baek EK, Kim YA, Yang SJ, Choi IY, et al. Efficacy of the smartphone-based glucose management application stratified by user satisfaction. Diabetes Metab J 2014 Jun;38(3):204-210 [FREE Full text] [doi: 10.4093/dmj.2014.38.3.204] [Medline: 25003074]
46. Klasnja P, Hartzler A, Powell C, Pratt W. Supporting cancer patients' unanchored health information management with mobile technology. AMIA Annu Symp Proc 2011;2011:732-741 [FREE Full text] [Medline: 22195130]
47. Liu W, Huang C, Wang C, Lee K, Lin S, Kuo H. A mobile telephone-based interactive self-care system improves asthma control. Eur Respir J 2011 Feb;37(2):310-317 [FREE Full text] [doi: 10.1183/09031936.00000810] [Medline: 20562122]

48. Logan AG, Irvine MJ, McIsaac WJ, Tisler A, Rossos PG, Easty A, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension 2012 Jul;60(1):51-57 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.188409] [Medline: 22615116]
49. Maguire R, Ream E, Richardson A, Connaghan J, Johnston B, Kotronoulas G, et al. Development of a novel remote patient monitoring system: the advanced symptom management system for radiotherapy to improve the symptom experience of patients with lung cancer receiving radiotherapy. Cancer Nurs 2015;38(2):E37-E47. [doi: 10.1097/NCC.0000000000000150] [Medline: 24836956]
50. Mira JJ, Navarro I, Botella F, Borrás F, Nuño-Solinís R, Orozco D, et al. A Spanish pillbox app for elderly patients taking multiple medications: randomized controlled trial. J Med Internet Res 2014;16(4):e99 [FREE Full text] [doi: 10.2196/jmir.3269] [Medline: 24705022]
51. Nes AAG, van Dulmen S, Eide E, Finset A, Kristjánsdóttir OB, Steen IS, et al. The development and feasibility of a web-based intervention with diaries and situational feedback via smartphone to support self-management in patients with diabetes type 2. Diabetes Res Clin Pract 2012 Sep;97(3):385-393. [doi: 10.1016/j.diabres.2012.04.019] [Medline: 22578890]
52. Patel S, Jacobus-Kantor L, Marshall L, Ritchie C, Kaplinski M, Khurana PS, et al. Mobilizing your medications: an automated medication reminder application for mobile phones and hypertension medication adherence in a high-risk urban population. J Diabetes Sci Technol 2013 May 01;7(3):630-639 [FREE Full text] [Medline: 23759395]
53. Pellegrini CA, Hoffman SA, Daly ER, Murillo M, Iakovlev G, Spring B. Acceptability of smartphone technology to interrupt sedentary time in adults with diabetes. Transl Behav Med 2015 Sep;5(3):307-314 [FREE Full text] [doi: 10.1007/s13142-015-0314-3] [Medline: 26327936]
54. Quinn CC, Khokhar B, Weed K, Barr E, Gruber-Baldini AL. Older adult self-efficacy study of mobile phone diabetes management. Diabetes Technol Ther 2015 Jul;17(7):455-461 [FREE Full text] [doi: 10.1089/dia.2014.0341] [Medline: 25692373]
55. Quinn CC, Shardell MD, Terrin ML, Barr EA, Park D, Shaikh F, et al. Mobile diabetes intervention for glycemic control in 45- to 64-year-old persons with type 2 diabetes. J Appl Gerontol 2016 Feb;35(2):227-243. [doi: 10.1177/0733464814542611] [Medline: 25098253]
56. 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]
57. Schreier G, Eckmann H, Hayn D, Kreiner K, Kastner P, Lovell N. Web versus app: compliance of patients in a telehealth diabetes management programme using two different technologies. J Telemed Telecare 2012 Dec;18(8):476-480. [doi: 10.1258/jtt.2012.GTH112] [Medline: 23209270]
58. Seto E, Leonard KJ, Masino C, Cafazzo JA, Barnsley J, Ross HJ. Attitudes of heart failure patients and health care providers towards mobile phone-based remote monitoring. J Med Internet Res 2010 Nov 29;12(4):e55 [FREE Full text] [doi: 10.2196/jmir.1627] [Medline: 21115435]
59. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012;14(1):e31 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 22356799]
60. Shinohara A, Ito T, Ura T, Nishiguchi S, Ito H, Yamada M, et al. Development of lifelog sharing system for rheumatoid arthritis patients using smartphone. Conf Proc IEEE Eng Med Biol Soc 2013;2013:7266-7269. [doi: 10.1109/EMBC.2013.6611235] [Medline: 24111422]
61. Stahlberg M, Govind S, Orr N, Gopal A, Lachmann J, Bhagirath R. Abstract 18895: Utility of remote monitoring using a mobile phone in patients with heart failure: the remote-HF-1 study. Circulation 2014 Nov 14;130(Suppl 2):A18895.
62. Tabak M, op den Akker H, Hermens H. Motivational cues as real-time feedback for changing daily activity behavior of patients with COPD. Patient Educ Couns 2014 Mar;94(3):372-378. [doi: 10.1016/j.pec.2013.10.014] [Medline: 24332934]
63. Tabak M, Vollenbroek-Hutten MM, van der Valk PD, van der Palen J, Hermens HJ. A telerehabilitation intervention for patients with chronic obstructive pulmonary disease: a randomized controlled pilot trial. Clin Rehabil 2014 Jun;28(6):582-591. [doi: 10.1177/0269215513512495] [Medline: 24293120]
64. Tatara N, Årsand E, Bratteteig T, Hartvigsen G. Usage and perceptions of a mobile self-management application for people with type 2 diabetes: qualitative study of a five-month trial. Stud Health Technol Inform 2013;192(1-2):127-131. [Medline: 23920529]
65. Triantafyllidis A, Velardo C, Chantler T, Shah SA, Paton C, Khorshidi R, SUPPORT-HF Investigators. A personalised mobile-based home monitoring system for heart failure: the SUPPORT-HF Study. Int J Med Inform 2015 Oct;84(10):743-753. [doi: 10.1016/j.ijmedinf.2015.05.003] [Medline: 26037921]
66. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. It's LiFe! Mobile and web-based monitoring and feedback tool embedded in primary care increases physical activity: a cluster randomized controlled trial. J Med Internet Res 2015 Jul 24;17(7):e184 [FREE Full text] [doi: 10.2196/jmir.4579] [Medline: 26209025]
67. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. A monitoring and feedback tool embedded in a counselling protocol to increase physical activity of patients with COPD or type 2 diabetes in primary

care: study protocol of a three-arm cluster randomised controlled trial. BMC Fam Pract 2014;15:93 [FREE Full text] [doi: 10.1186/1471-2296-15-93] [Medline: 24885096]
68. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. A pilot study of a tool to stimulate physical activity in patients with COPD or type 2 diabetes in primary care. J Telemed Telecare 2014 Jan;20(1):29-34. [doi: 10.1177/1357633X13519057] [Medline: 24414397]
69. Vuorinen A, Leppänen J, Kaijanranta H, Kulju M, Heliö T, van Gils M, et al. Use of home telemonitoring to support multidisciplinary care of heart failure patients in Finland: randomized controlled trial. J Med Internet Res 2014;16(12):e282 [FREE Full text] [doi: 10.2196/jmir.3651] [Medline: 25498992]
70. Waki K, Aizawa K, Kato S, Fujita H, Lee H, Kobayashi H, et al. DialBetics with a multimedia food recording tool, foodlog: smartphone-based self-management for type 2 diabetes. J Diabetes Sci Technol 2015 May;9(3):534-540 [FREE Full text] [doi: 10.1177/1932296815579690] [Medline: 25883164]
71. Waki K, Fujita H, Uchimura Y, Omae K, Aramaki E, Kato S, et al. DialBetics: a novel smartphone-based self-management support system for type 2 diabetes patients. J Diabetes Sci Technol 2014 Mar 13;8(2):209-215 [FREE Full text] [doi: 10.1177/1932296814526495] [Medline: 24876569]
72. Wayne N, Ritvo P. Smartphone-enabled health coach intervention for people with diabetes from a modest socioeconomic strata community: single-arm longitudinal feasibility study. J Med Internet Res 2014;16(6):e149 [FREE Full text] [doi: 10.2196/jmir.3180] [Medline: 24907918]
73. Williams V, Price J, Hardinge M, Tarassenko L, Farmer A. Using a mobile health application to support self-management in COPD: a qualitative study. Br J Gen Pract 2014 Jul;64(624):e392-e400 [FREE Full text] [doi: 10.3399/bjgp14X680473] [Medline: 24982491]
74. Wolf A, Olsson L, Swedberg K, Ekman I. Use of smartphones in person-centred eHealth diaries in patients above 75 years. Eur J Cardiovasc Nurs 2012;11(Suppl 1):P100.
75. Wood FG, Alley E, Baer S, Johnson R. Interactive multimedia tailored to improve diabetes self-management. Nurs Clin North Am 2015 Sep;50(3):565-576. [doi: 10.1016/j.cnur.2015.05.009] [Medline: 26333610]
76. Wu J, Yu H, Ho T, Su X, Lin M, Lai F. Tablet PC-enabled application intervention for patients with gastric cancer undergoing gastrectomy. Comput Methods Programs Biomed 2015 Apr;119(2):101-109. [doi: 10.1016/j.cmpb.2015.03.004] [Medline: 25819034]
77. Chow E, Faye L, Sawatzky R, Webster G, Cho MZ. PROMs background document. Ottawa, ON: Canadian Institute for Health Information; 2015. URL: https://www.cihi.ca/sites/default/files/document/proms_background_may21_en-web.pdf [accessed 2017-12-07] [WebCite Cache ID 6maQHvWh0]
78. Smith A. Older adults and technology use. Washington, DC: Pew Research Center; 2014 Apr 03. URL: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/ [accessed 2016-12-07] [WebCite Cache ID 6P34DTel4]
79. Ohlund B, Yu C. Threats to validity of research design. Portland, OR: Portland State University; 2003. URL: http://web.pdx.edu/~stipakb/download/PA555/ResearchDesign.html [accessed 2016-12-07] [WebCite Cache ID 6maQpSHZa]
80. Krishna S, Boren SA. Diabetes self-management care via cell phone: a systematic review. J Diabetes Sci Technol 2008 May;2(3):509-517 [FREE Full text] [Medline: 19885219]
81. Newman S, Steed L, Mulligan K. Self-management interventions for chronic illness. Lancet 2004;364(9444):1523-1537. [doi: 10.1016/S0140-6736(04)17277-2] [Medline: 15500899]
82. Thompson WG, Kuhle CL, Koepp GA, McCrady-Spitzer SK, Levine JA. "Go4Life" exercise counseling, accelerometer feedback, and activity levels in older people. Arch Gerontol Geriatr 2014;58(3):314-319. [doi: 10.1016/j.archger.2014.01.004] [Medline: 24485546]
83. Marks R, Allegrante JP, Lorig K. A review and synthesis of research evidence for self-efficacy-enhancing interventions for reducing chronic disability: implications for health education practice (part II). Health Promot Pract 2005 Apr;6(2):148-156. [doi: 10.1177/1524839904266792] [Medline: 15855284]
84. DeWalt DA, Davis TC, Wallace AS, Seligman HK, Bryant-Shilliday B, Arnold CL, et al. Goal setting in diabetes self-management: taking the baby steps to success. Patient Educ Couns 2009 Nov;77(2):218-223 [FREE Full text] [doi: 10.1016/j.pec.2009.03.012] [Medline: 19359123]
85. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. N Engl J Med 2012 Mar 01;366(9):780-781. [doi: 10.1056/NEJMp1109283] [Medline: 22375967]
86. Huang A, Chen C, Bian K, Duan X, Chen M, Gao H, et al. WE-CARE: an intelligent mobile telecardiology system to enable mHealth applications. IEEE J Biomed Health Inform 2014 Mar;18(2):693-702. [doi: 10.1109/JBHI.2013.2279136] [Medline: 24608067]
87. Chan M, Estève D, Fourniols J, Escriba C, Campo E. Smart wearable systems: current status and future challenges. Artif Intell Med 2012 Nov;56(3):137-156. [doi: 10.1016/j.artmed.2012.09.003] [Medline: 23122689]
88. Dailey J, Kalis B, Sather T. Insight driven health: silver surfers continue to make waves in ehealth. But are healthcare organizations swimming against the tide?. Chicago, IL: Accenture; 2016. URL: https://www.accenture.com/t20160829T222305__w__/us-en/_acnmedia/PDF-29/Accenture-Silver-Surfers-Continue-to-Make-Waves.pdf [accessed 2017-03-30] [WebCite Cache ID 6pMLflNvl]

89. 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;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
90. Bender MS, Santos G, Villanueva C, Arai S. Development of a mobile phone-based weight loss lifestyle intervention for Filipino Americans with type 2 diabetes: protocol and early results from the PilAm Go4Health randomized controlled trial. JMIR Res Protoc 2016 Sep 08;5(3):e178 [FREE Full text] [doi: 10.2196/resprot.5836] [Medline: 27608829]
91. Azar KMJ, Koliwad S, Poon T, Xiao L, Lv N, Griggs R, et al. The Electronic CardioMetabolic Program (eCMP) for patients with cardiometabolic risk: a randomized controlled trial. J Med Internet Res 2016 May 27;18(5):e134 [FREE Full text] [doi: 10.2196/jmir.5143] [Medline: 27234480]
92. Oresko JJ, Duschl H, Cheng AC. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. IEEE Trans Inf Technol Biomed 2010 May;14(3):734-740. [doi: 10.1109/TITB.2010.2047865] [Medline: 20388600]
93. Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA 2013 Apr 3;309(13):1351-1352. [doi: 10.1001/jama.2013.393] [Medline: 23549579]
94. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2:3 [FREE Full text] [doi: 10.1186/2047-2501-2-3] [Medline: 25825667]
95. Zacharaki EI, Wang S, Chawla S, Soo YD, Wolf R, Melhem ER, et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med 2009 Dec;62(6):1609-1618 [FREE Full text] [doi: 10.1002/mrm.22147] [Medline: 19859947]
96. Sharafoddini A, Dubin JA, Lee J. Patient similarity in prediction models based on health data: a scoping review. JMIR Med Inform 2017 Mar 03;5(1):e7 [FREE Full text] [doi: 10.2196/medinform.6730] [Medline: 28258046]
97. Rabbi M, Pfammatter A, Zhang M, Spring B, Choudhury T. Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. JMIR Mhealth Uhealth 2015;3(2):e42 [FREE Full text] [doi: 10.2196/mhealth.4160] [Medline: 25977197]
98. Giansanti D, Maccioni G, Cesinaro S, Benvenuti F, Macellari V. Assessment of fall-risk by means of a neural network based on parameters assessed by a wearable device during posturography. Med Eng Phys 2008 Apr;30(3):367-372. [doi: 10.1016/j.medengphy.2007.04.006] [Medline: 17560825]
99. Bellos CC, Papadopoulos A, Rosso R, Fotiadis DI. Identification of COPD patients' health status using an intelligent system in the CHRONIOUS wearable platform. IEEE J Biomed Health Inform 2014 May;18(3):731-738. [doi: 10.1109/JBHI.2013.2293172] [Medline: 24808219]
100. Ferrante do Amaral CE, Wolf B. Current development in non-invasive glucose monitoring. Med Eng Phys 2008 Jun;30(5):541-549. [doi: 10.1016/j.medengphy.2007.06.003] [Medline: 17942360]
101. Vaddiraju S, Burgess DJ, Tomazos I, Jain FC, Papadimitrakopoulos F. Technologies for continuous glucose monitoring: current problems and future promises. J Diabetes Sci Technol 2010 Nov 01;4(6):1540-1562 [FREE Full text] [doi: 10.1177/193229681000400632] [Medline: 21129353]

Abbreviations

CCM: chronic care model
COPD: chronic obstructive pulmonary disease
DSS: decision support system
EQ-5D: EuroQoL Five Dimensions Questionnaire
HbA1c: hemoglobin A1c
PROMs: patient-reported outcome measures
RCT: randomized controlled trial
SF-36: 36-Item Short Form Health Survey
SMS: short message service
T2DM: type 2 diabetes mellitus

Edited by P Bamidis; submitted 10.12.16; peer-reviewed by A Triantafyllidis, N Tatara, P Toumpaniaris; comments to author 23.02.17; revised version received 30.03.17; accepted 18.04.17; published 23.05.17.

Please cite as:
Kim BYB, Lee J
Smart Devices for Older Adults Managing Chronic Disease: A Scoping Review
JMIR Mhealth Uhealth 2017;5(5):e69
URL: http://mhealth.jmir.org/2017/5/e69/
doi:10.2196/mhealth.7141
PMID:28536089

Original Paper

Cognitive Testing in People at Increased Risk of Dementia Using a Smartphone App: The iVitality Proof-of-Principle Study

Susan Jongstra[1], MD; Liselotte Willemijn Wijsman[2], PhD, MD; Ricardo Cachucho[3], MSc; Marieke Peternella Hoevenaar-Blom[1], MSc, PhD; Simon Pieter Mooijaart[2,4], MD, PhD; Edo Richard[1,5], MD, PhD

[1]Academic Medical Center, Department of Neurology, University of Amsterdam, Amsterdam, Netherlands

[2]Leiden University Medical Center, Department of gerontology & geriatrics, University of Leiden, Leiden, Netherlands

[3]Leiden Institute of Advanced Computer Sciences, LIACS, University of Leiden, Leiden, Netherlands

[4]Institute of Evidence-based Medicine in Old Age - IEMO, LUMC, Leiden, Netherlands

[5]Donders Institute for Brain, Behaviour and Cognition, Department of Neurology, Radboud UMC, Nijmegen, Netherlands

Corresponding Author:
Susan Jongstra, MD
Academic Medical Center
Department of Neurology
University of Amsterdam
Meiberdreef 9
Amsterdam, 1105 AZ
Netherlands
Phone: 31 205663446
Fax: 31 205669290
Email: s.jongstra@amc.uva.nl

Abstract

Background: Smartphone-assisted technologies potentially provide the opportunity for large-scale, long-term, repeated monitoring of cognitive functioning at home.

Objective: The aim of this proof-of-principle study was to evaluate the feasibility and validity of performing cognitive tests in people at increased risk of dementia using smartphone-based technology during a 6 months follow-up period.

Methods: We used the smartphone-based app iVitality to evaluate five cognitive tests based on conventional neuropsychological tests (Memory-Word, Trail Making, Stroop, Reaction Time, and Letter-N-Back) in healthy adults. Feasibility was tested by studying adherence of all participants to perform smartphone-based cognitive tests. Validity was studied by assessing the correlation between conventional neuropsychological tests and smartphone-based cognitive tests and by studying the effect of repeated testing.

Results: We included 151 participants (mean age in years=57.3, standard deviation=5.3). Mean adherence to assigned smartphone tests during 6 months was 60% (SD 24.7). There was moderate correlation between the firstly made smartphone-based test and the conventional test for the Stroop test and the Trail Making test with Spearman ρ=.3-.5 (P<.001). Correlation increased for both tests when comparing the conventional test with the mean score of all attempts a participant had made, with the highest correlation for Stroop panel 3 (ρ=.62, P<.001). Performance on the Stroop and the Trail Making tests improved over time suggesting a learning effect, but the scores on the Letter-N-back, the Memory-Word, and the Reaction Time tests remained stable.

Conclusions: Repeated smartphone-assisted cognitive testing is feasible with reasonable adherence and moderate relative validity for the Stroop and the Trail Making tests compared with conventional neuropsychological tests. Smartphone-based cognitive testing seems promising for large-scale data-collection in population studies.

(JMIR Mhealth Uhealth 2017;5(5):e68) doi:10.2196/mhealth.6939

KEYWORDS

telemedicine; cognition; neuropsychological tests

Introduction

The global prevalence of dementia is likely to increase in the coming years, mainly due to the growing population with an increased life expectancy [1]. To investigate interventions to prevent dementia, large sample sizes with long follow-up are required [2]. Assessment of cognitive functioning over time is important for early detection of cognitive decline in longitudinal dementia prevention studies. Conventional neuropsychological examination is burdensome, time-consuming, and expensive and therefore hardly feasible in large-scale studies with long follow-up. To get informed about cognitive functioning without the need for full neuropsychological examination, innovative solutions are required.

New technology is rapidly adopted by older generations, illustrated by a steady increase in the Internet and smartphone use over the last years [3]. Using smartphone technology, remote monitoring of health parameters such as physical activity and blood pressure have already been widely studied and found feasible, also in older populations [4,5]. Smartphones are likely to be the principal platform for the development of the next generation of clinical care and research [6]. Smartphone-assisted cognitive testing would provide the ability to assess cognitive functioning rapidly and repeatedly in a noninvasive manner, at a convenient moment, and without generating high costs. Experience with smartphone use during a clinical cognitive assessment has already been tested [7], paving the way to integration in a home setting. Feasibility and validity of smartphone-based cognitive testing has been described, although narrowed down to specific patient groups or a specific cognitive test [8-11]. Despite these advances made in conducting smartphone research, little is known in terms of the feasibility and validity of applying multiple cognitive tests using smartphone-based technologies for clinical research in larger populations. Implementation of an app is only feasible if participants are compliant [12] and the technical performance is optimal [13].

The aim of this study was to investigate the feasibility and validity of a cognitive test battery using smartphone-assisted technology in healthy adults, during a 6 months follow-up period.

Methods

Study Participants

We recruited participants at increased risk of cognitive decline and dementia, operationalized as a parental history of dementia [14]. These persons are highly motivated to participate in a monitoring study to support preventive strategies for dementia, and therefore suitable for a proof-of-principle study [15].

Participants were included if: (1) they were 50 years or older, (2) at least one of their parents was diagnosed with any form of dementia, (3) they knew how to handle and were in possession of a smartphone with iOS or Android (version 2.3.3 or higher) software, (4) they had no dementia or any other cognitive disorder, and (5) they had no medical history of stroke or transient ischemic attack.

Participants were recruited through advertisements at memory outpatient clinics, nursing homes, general practices, and using the communication channels (website and newsletter) of the Dutch Alzheimer Foundation. People were asked to contact the study center and if all of the inclusion criteria were met, participants received detailed study information in print and an appointment for baseline measurement was made. Enrolment and follow-up took place from September 2013 to January 2015. Written informed consent was obtained from all participants at the baseline study visit. The study was approved by the medical ethical committee of Leiden University Medical Center (LUMC), the Netherlands.

iVitality Platform

iVitality is a Web-based research platform that consists of a website, a smartphone-based app, and sensors that are connected with or already integrated in the smartphone to measure health characteristics including cognitive function, blood pressure [4], physical activity (integrated pedometer), and life style (with questions about health and mood). The smartphone-based app was installed during the baseline assessment and the sensors were activated if participants were officially included in the study, until the end of follow-up. Participants could log on to the website to overlook the measurements and results of their performance on the app. Participants received alerts from the iVitality smartphone app to perform a test or measurement (eg, cognitive test or blood pressure) on their smartphone.

Study Design

Participants visited the study center at LUMC or Academic Medical Center (AMC) at baseline, where they received information about the study and the smartphone-based app was installed and explained. During this visit, baseline measurements were performed by a study physician or research nurse. Afterwards, during a 6 month follow-up period, participants received messages on their smartphone, reminding them to voluntarily perform a specific cognitive test (Table 1). Alert moments were chosen in a way that every test had at least four reminder moments evenly spread during the 6 month follow-up period. Table 1 indicates on what day since baseline the message was sent for every test to every individual participant. The smartphone app collected data from the tests and provided feedback to the participant by showing the results of their measurements. A secured Internet connection transferred the data to the website and the database of the study center.

Table 1. Message moment per cognitive test during follow-up.

Weeks in study	1	3	5	7	9	11	13	15	17	19	21	23	25
Memory-Word	Day 1		Day 29					Day 99					Day 169
Trail Making test	Day 2			Day 43					Day113				Day 170
Stroop	Day 3				Day57					Day127			Day 171
Reaction Time test	Day 4					Day71					Day 141		Day 172
Letter-N-Back		Day 15					Day 85					Day 155	Day 173

Baseline Measurements

In preparation for the first visit to the study center, all participants completed a Web-based questionnaire including questions about level of education, medical history, and medication use. The study physician measured parameters including weight, height, and blood pressure of all participants.

Cognitive function at baseline was tested using five neuropsychological tests to assess global cognitive function, executive function, attention, and immediate and delayed recall. The mini-mental state examination (MMSE) [16] was used to evaluate global cognitive function. The 15-Word Verbal Learning test (15-WVLT) [17] was used to assess immediate and delayed recall. The Trail Making test (TMT) [18], parts A and B, were used to measure attention and executive function. The Stroop-Color-Word test [19] was used to test selective attention.

Smartphone-Based Cognitive Tests

Five digital versions of cognitive tests were developed for the iVitality smartphone app based on existing neuropsychological tests, but carefully adapted for smartphone use.

The Memory-Word test was based on the 15-WVLT [17]. A series of 10 words with a fixed time pace was presented to the participants, which they were instructed to remember. Directly afterwards, participants were displayed a list of 20 words, including the 10 words which were presented before, mixed with 10 new words. Participants had to press “yes” or “no” for recognition. Each correct and incorrect response was recorded.

The TMT, based on the original TMT part A and B [18], consisted of four parts of increasing complexity in which participants had to make a trail connecting 12 circles. In part 1, the circles contained numbers in ascending order (1-2-3), part 2 contained letters in ascending order (A-B-C), part 3 contained numbers and letters alternating in ascending order (1-A-2-B), and in part 4, numbers and letters had to be connected alternating and in opposing order: numbers ascending, letters descending (1-Z-2-Y). The total time for each part was recorded. This last part was added to decrease the ceiling effect in a cognitively healthy population.

The Stroop color-word test was based on the original Stroop test [19]. In the smartphone version, 30 items were presented in all three parts. Names of colors in black letters (part I), colored blocks (part II) or names of colors in other colored letters (part III) were presented together with multiple-choice answers. Total time to complete each part was recorded.

The Reaction Time test consisted of two parts: in part 1, participants were requested to touch the screen of the smartphone as soon as a presented green box turned blue. In part 2, the green box was again presented, but turned into either a blue or red box. The participants had to touch the screen as soon as possible, only if the blue box appeared. At one random instance an enlarged blue box was presented, as a measure of the startle time. In all parts, the time was recorded between the box turning blue and the moment the participant touched the screen in milliseconds. The time between presenting the enlarged blue box and pressing the screen was recorded as the startle (reaction) time in milliseconds.

The Letter-N-Back test, based on the original N-back test [20], consisted of four parts. A series of letters on the screen of the smartphone was presented in a sequential order. In part 1 (0-back), participants had to touch the screen when the letter “X” appeared (in total 11 items presented); for part 2 (1-back), participants had to touch the screen when the letter that was displayed, was the same as the previous one (in total 11 items presented); in part 3 (2-back), participants had to touch the screen when a letter that was displayed was the same as the one before the previous one (in total 15 items presented); and in part 4 (3-back), they had to touch the screen when the letter that was displayed was the same as the one that was presented before the previous 2 letters (in total 20 items presented). Each correct and incorrect response was recorded.

Prior to each test, a short explanation was displayed. Screenshots of the tests are shown in Multimedia Appendix 1.

Statistical Analyses

Characteristics of the study participants are reported as mean (SD) for continuous variables and as number (%) for categorical variables.

Feasibility was evaluated by the technical performance of the app and adherence to perform cognitive tests on a smartphone. Validity was studied by assessing the correlation between conventional and smartphone cognitive tests, and the effect on performance of repeated cognitive tests on a smartphone.

For each participant and each test, we assessed adherence during follow-up. Adherence was defined as the actual performance of cognitive test measurements within 1 week of the reminder received through the smartphone app. The technical performance was defined as the ability to function as developed on every participant’s smartphone.

To assess the relative validity of the first performed smartphone test compared with the conventional Stroop and TMT, we

calculated the correlation coefficient. Since the test results were generally not normally distributed, we used Spearman correlation coefficient. To investigate systematic differences between conventional and smartphone cognitive tests, we computed z-scores for both and visualized the values in a Bland-Altman plot.

In a sensitivity analysis, we assessed the correlation between the score on the conventional test at baseline and the mean score of all attempts a participant had made on a specific smartphone-based test, to account for (technical) difficulties in the first attempt and a learning curve. In a second sensitivity analysis, we assessed the correlation between the conventional Stroop test and the first smartphone attempt without many mistakes. The participant needed to score at least half of the answers correct, and if not, the following score (of the next attempt) was taken. Since no conventional version of the Letter-N-Back test and the Reaction Time test were done at baseline, we could not assess the relative validity for these tests.

To assess potential learning effects after repeated testing, performance over time on the smartphone cognitive tests were visualized graphically. We analyzed the linear trend in test performance with each attempt using a linear mixed effects model with a random intercept and random slope for attempt within each subject (MIXED procedure). To investigate selective dropout, we performed an additional analysis on the effect of repeated testing including only those participants who performed 9 tests or more.

All analyses were performed using IBM SPSS software (version 23).

Results

Baseline Characteristics

The flowchart for inclusion of participants is shown in Figure 1. The study population consisted of 151 participants. Two participants discontinued the study immediately after baseline visit because of technical issues with their smartphone, so they do not have smartphone measurements. During the follow-up period of 6 months, 12 participants (8.1%, 12/149) discontinued the study.

Baseline characteristics are shown in Table 2. Mean age was 57.3 (SD 5.3) years and 70.9% (107/151) were female. The most commonly used smartphone types were iPhone and Samsung. Around 58.3% (88/151) of the participants had a high education level. None of the participants had an MMSE score below 27 points. More details about the other baseline characteristics are published elsewhere [4,15].

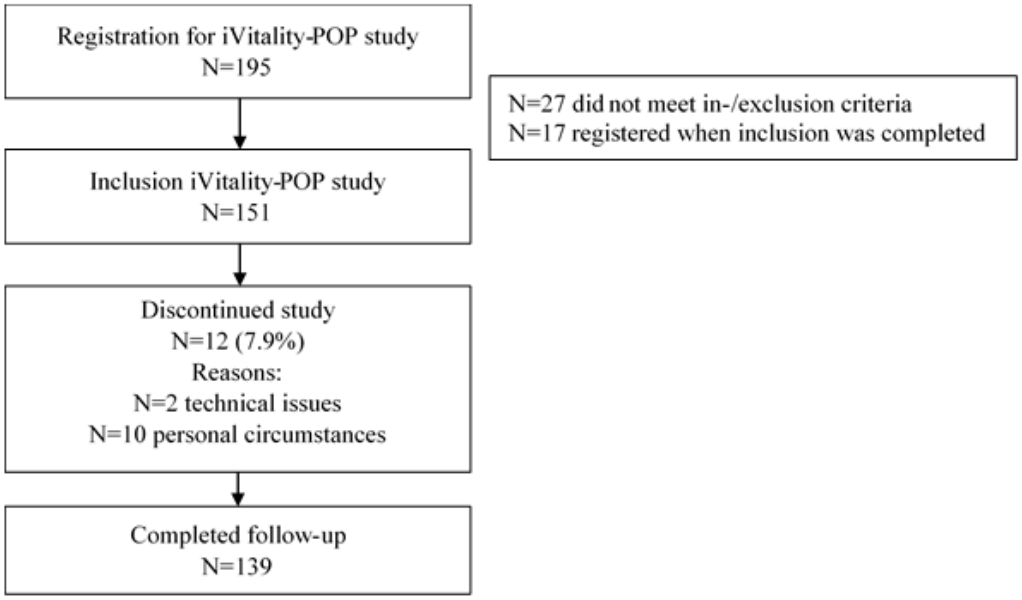
Table 2. Baseline characteristics of study participants.

Demographics	Study participants (N=151)
Age (years), mean (SD)	57.3 (5.3)
Female, n (%)	107 (70.9)
Highest education level[a] , n (%)	
Low (<7 years)	16 (10.6)
Middle (7-12 years)	44 (29.1)
High (>12 years)	88 (58.3)
Body mass index (kg/m^2), mean (SD)	26.4 (4.0)
Systolic blood pressure (mmHg), mean (SD)	138 (18.2)
Diastolic blood pressure (mmHg), mean (SD)	85 (10.8)
MMSE[b], median (interquartile range)	29 (29-30)

[a]Missing data for n=3 participants.

[b]MMSE: mini mental state examination.

Figure 1. Flowchart inclusion of study participants.



Adherence

Adherence to the test program of the five smartphone-based cognitive tests during a 6 month follow-up is shown in Figure 2. Adherence was highest for the Reaction Time test (67%) and slightly lower for the other tests (62% for the Stroop test, 61% for the Memory-Word test, 61% for the Letter-N-Back test, and 48% for the Trail Making test). During the 6 month follow-up, adherence slightly decreased for all tests. Mean adherence per participant was 60% (SD 24.7). When investigating the data for the percentage of participants (calculated from total N=151) who made a test at least once during follow-up, irrespective of timing relative to the reminder, this was 98% for the Reaction Time test, 97% for the Stroop test, 95% for the N-back test, 94% for the Memory-Word test, and 89% for the TMT.

Relative Validity of the Smartphone Test Compared With the Conventional Test

Raw test scores of the conventional tests at baseline and the firstly performed smartphone tests are described in Multimedia Appendix 2. Since the smartphone-based tests were based on the conventional tests but not identical, direct comparison between the raw test scores is not possible using absolute values.

Figure 2. Adherence to smartphone-based cognitive tests.

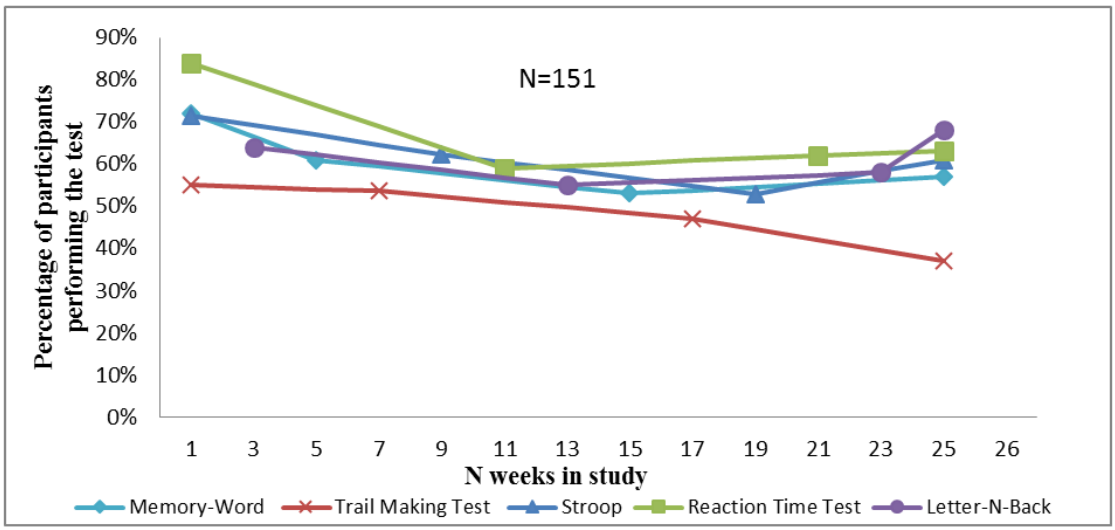


Table 3. Relative validity of conventional cognitive test at baseline compared with cognitive test on smartphone.

Test	n	Conventional versus first performed smartphone test		Conventional versus mean score of all performed smartphone tests	
		ρ CC[a]	*P* value	ρ CC	*P* value
Stroop panel 1	146	.36	<.001[b]	.42	<.001[b]
Stroop panel 2	146	.31	<.001[b]	.36	<.001[b]
Stroop panel 3	146	.50	<.001[b]	.62	<.001[b]
TMT[c] numeric	135	.38	<.001[b]	.48	<.001[b]
TMT alphanumeric	135	.39	<.001[b]	.43	<.001[b]

[a]CC: correlation coefficient.

[b]Significant at the <.001 level.

[c]TMT: Trail Making test.

The association between the conventional cognitive test made at baseline and the corresponding firstly performed cognitive test on the smartphone is shown in Table 3. There was moderate correlation between the smartphone-based test and the conventional test for the Stroop test (panel 3) and the TMT with ρ=.5 and .4 respectively.

The sensitivity analysis in which we investigated the correlation between the conventional test and the mean score of all performed corresponding smartphone tests during follow-up showed higher correlation coefficients for both tests compared with the correlation with the first performed cognitive test (Table 3).

The number of mistakes made by the participants in the conventional Stroop test was very low and randomly distributed, and therefore not accounted for in the analysis. The number of mistakes in the smartphone-based Stroop test was accounted for in a sensitivity analysis (Multimedia Appendix 3). This showed a higher correlation coefficient for all three panels compared with the correlation with the first performed cognitive test when not accounted for mistakes (panel 1: ρ=.39, panel 2: ρ=.33, and panel 3: ρ=.57).

The Bland-Altman plots of tests which showed moderate correlation (Figure 3) show that the difference between the measurements was randomly distributed over the mean of the measurements. However, inspection of the Bland-Altman plot suggests that for the TMT (numeric and alphanumeric), correlation decreases with increasing time needed to complete the test.

Repeated Cognitive Testing

The trend in test scores for each smartphone-based test is shown in Figure 4. With increasing number of test repeats, the number of participants contributing to the data decreased since each test was actively offered 4 times during the study, so any excess number of performed tests is on participants' initiative. The performance on the Stroop test improved for each panel with almost 1 sec per attempt (panel 3: beta=−.93, *P*<.001) and the reversed alphanumeric TMT improved with 1.8 sec per attempt (beta=−1.81, *P*<.001). The performance on the N-back, the Memory-Word, and the Reaction Time test remained virtually stable over time.

The sensitivity analysis in participants who performed the tests at least nine times showed similar results (Multimedia Appendix 4).

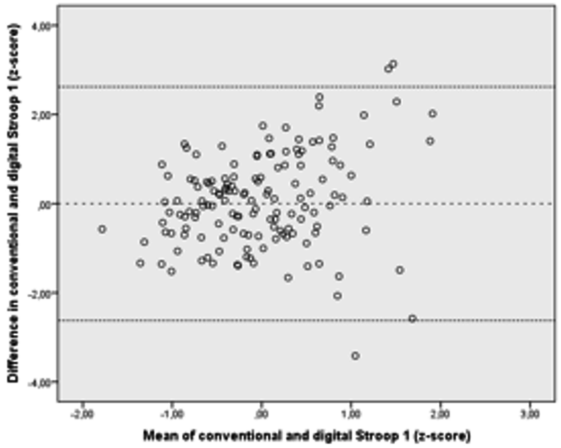
Discussion

Principal Findings

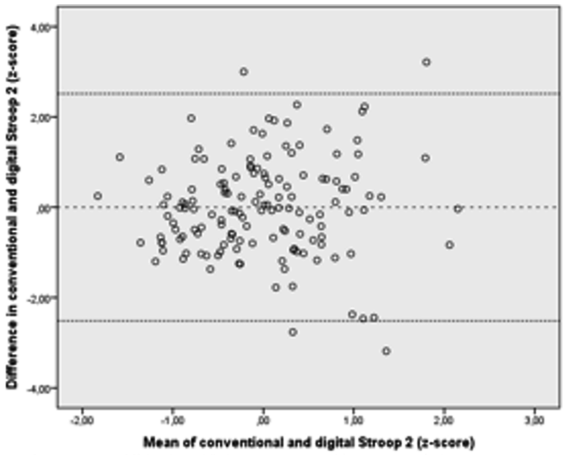
Our study shows that smartphone-based cognitive testing in cognitively healthy adults over 50 years of age is feasible and that motivated research participants are reasonably adherent to regular testing following an alert on their smartphone. Of the cognitive tests developed in iVitality, the smartphone-based Stroop test and the TMT have a moderate correlation with conventional tests. Repeated testing leads to improved test scores with increasing number of tests performed, suggesting a learning effect for the Stroop test and the TMT.

Adherence to smartphone tests in trial setting varies between studies (17%-90%) [21-23]. These mixed percentages can be explained by the broad definition of adherence in smartphone interventions considering different frequencies, lengths, and intensities of use. Adherence of our participants is relatively good (60%) compared with these studies. The high frequency of reminders the participants received not only for the cognitive tests, but also for the other measurements in the iVitality POP study, could have caused a certain degree of alarm fatigue. This could have reduced the adherence and might explain the variability in adherence in our study. Participants were most adherent to the Reaction Time test. Potential reasons are that the test is easy, not very time-consuming, and does not require processing of information. Only 2 participants (1.3%) could not perform the smartphone tests because of technical problems. This suggests that repeated smartphone-based neuropsychological testing outside the context of a research center is also technically feasible.

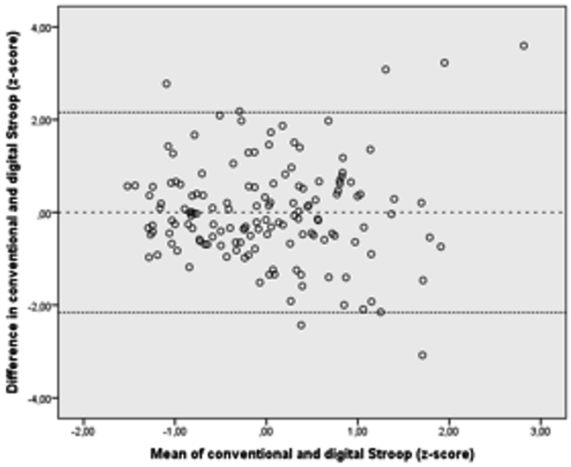
Figure 3. Systematic differences between conventional and smartphone-based cognitive tests in a Bland-Altman-plot. All values are standardized in z-scores.



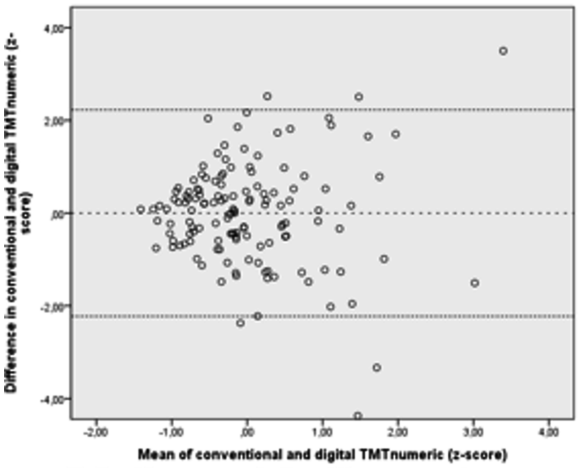
Stroop panel 1. Dotted lines represent the mean difference (0,0), +2SD (2,30) and -2SD (-2,30)



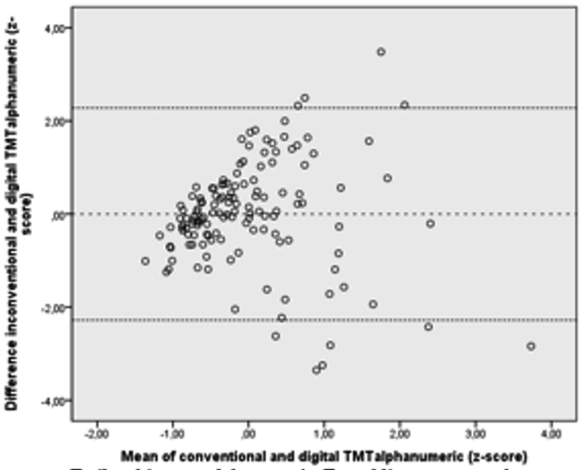
Stroop panel 2. Dotted lines represent the mean difference (0,0), +2SD (2,51) and -2SD (-2,51)



Stroop panel 3. Dotted lines represent the mean difference (0,0), +2SD (2,16) and -2SD (-2,16)

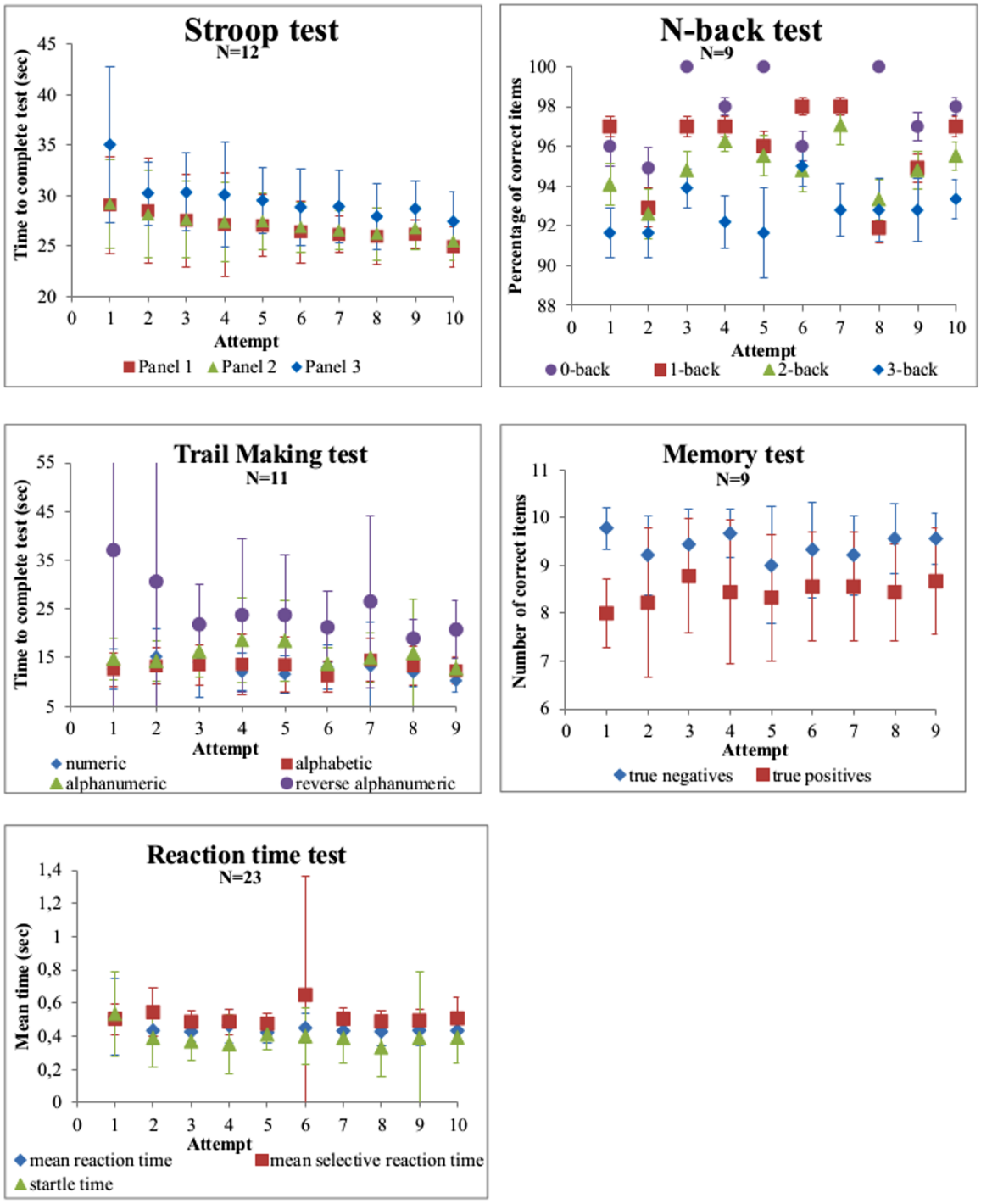


Trail making test numeric. Dotted lines represent the mean difference (0,0), +2SD (2,23) and -2SD(-2,23)



Trail making test alphanumeric. Dotted lines represent the mean difference (0,0), +2SD (2,21) and -2SD (-2,21)

Figure 4. Effect of repeated cognitive tests on smartphone.



Few studies have been performed to validate cognitive testing using a smartphone, usually in the context of a specific disorder or healthy young people [8,10,22]. The moderate correlations in our study for the attention and executive function tasks are comparable with correlations found in a other study investigating cognitive smartphone apps focusing on working memory and perceptual speed [24]. Another Stroop smartphone app was already validated to diagnose covert hepatic encephalopathy [9], but was not compared with the conventional Stroop test [19]. The moderate association found between the conventional Stroop test and a smartphone Stroop test was not found before [22]. This is also the first study investigating the TMT on a smartphone compared with the conventional version [25] with a moderate correlation.

The correlation coefficient increases for all smartphone-based tests with more attempts and when leaving the scores out from participants who made many mistakes in the smartphone Stroop test (Multimedia Appendix 3), implicating that technical

challenges while performing the test may have to be overcome. Our participants received short digital instructions prior to the smartphone-based tests in an attempt to limit the influence of technical issues. Nevertheless, the first attempt could be less reliable because of misunderstanding. The mean performance reduced random measurement error and therefore resulted in stronger associations. Especially for the Stroop test we noticed that some participants made a lot of mistakes in the first attempt (more than half of the answers were incorrect), indicating misunderstanding and implicating the need for more explanation on beforehand in further research.

In line with our findings, another study that also developed a Letter-N-Back test and Reaction Speed test for the smartphone did not observe a learning effect over time [22]. The fact that we did not find an improvement on performance in the Memory-Word, Letter-N-Back, and Reaction Speed tests can be due to the ceiling effect in our sample of participants without any cognitive complaints.

Limitations

This proof-of-principle study has several limitations. We selected participants with a parental history of dementia and therefore they are highly motivated to participate. This may have introduced a selection bias toward better adherence, which reduces the external validity. Another limitation is that we could not validate every smartphone test to a conventional test administered at baseline. Future studies must try to develop a more comparable smartphone test. Strengths of this study are the relatively large sample size for a proof-of-principle study, the moderate level of adherence, and the validation of part of the tests to conventional neuropsychological tests.

Conclusions

Taken together, the results of this proof-of-principle study show that smartphone cognitive testing in healthy older individuals is feasible and yields valid test results. It allows for repeated testing to observe changes over time while reducing the need for face-to-face contact, making it time-efficient, less burdensome for research participants, and less expensive. The tests should be considered as screening tests to detect changes over time, rather than replacing conventional neuropsychological test batteries. It may be particularly useful for large-scale data-collection in population studies with long follow-up requiring repeated testing.

Before implementation of this type of tests, further research should focus on criterion validity to investigate whether the tests adequately pick up cognitive decline both cross-sectionally as well as longitudinally. To reduce a potential learning effect, alternative versions of the tests could be developed, although for longitudinal research this is less important since the learning effect seemed to wane in our study.

Acknowledgments

The authors thank Dr AJM de Craen, clinical epidemiologist, Dr BA Schmand, professor of neuropsychology, Dr WA van Gool, professor of neurology, and Dr EP Moll van Charante, general practitioner, for their comments on the analysis plan and interpretation of the data.

This project was supported by the Dutch Ministry of Health, Welfare and Sport and was enabled by ZonMw (NGI/NWO 050-060-810 project 627002001). The authors Susan Jongstra and Edo Richard received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 305374.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of cognitive tests on smartphone.

[PNG File, 261KB - mhealth_v5i5e68_app1.png]

Multimedia Appendix 2

Mean test results of conventional cognitive tests at baseline and first made cognitive test on a smartphone.

[PNG File, 1MB - mhealth_v5i5e68_app2.png]

Multimedia Appendix 3

Relative validity of the first performed cognitive test without many mistakes compared to the conventional cognitive test at baseline.

[PNG File, 298KB - mhealth_v5i5e68_app3.png]

Multimedia Appendix 4

Effect over time for repeated cognitive tests (only for the participants who made the test at least 9 times).

[PNG File, 588KB - mhealth_v5i5e68_app4.png]

References

1. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. Alzheimers Dement 2013 Jan;9(1):63-75.e2. [doi: 10.1016/j.jalz.2012.11.007] [Medline: 23305823]
2. Richard E, Andrieu S, Solomon A, Mangialasche F, Ahtiluoto S, Moll van Charante EP, et al. Methodological challenges in designing dementia prevention trials - the European Dementia Prevention Initiative (EDPI). J Neurol Sci 2012 Nov 15;322(1-2):64-70. [doi: 10.1016/j.jns.2012.06.012] [Medline: 22818266]
3. CentraalBureauStatistiek. 2016. Aandeel internetgebruikers met dagelijks internetgebruik URL: https://www.cbs.nl/nl-nl/nieuws/2016/22/acht-procent-van-de-nederlanders-nooit-op-internet [accessed 2017-04-26] [WebCite Cache ID 6q14BbVc2]
4. Wijsman LW, Richard E, Cachucho R, de Craen AJ, Jongstra S, Mooijaart SP. Evaluation of the use of home blood pressure measurement using mobile phone-assisted technology: the iVitality proof-of-principle study. JMIR Mhealth Uhealth 2016 Jun 13;4(2):e67 [FREE Full text] [doi: 10.2196/mhealth.5485] [Medline: 27296628]
5. Fong SS, Ng SS, Cheng YT, Zhang J, Chung LM, Chow GC, et al. Comparison between smartphone pedometer applications and traditional pedometers for improving physical activity and body mass index in community-dwelling older adults. J Phys Ther Sci 2016 May;28(5):1651-1656 [FREE Full text] [doi: 10.1589/jpts.28.1651] [Medline: 27313391]
6. Dufau S, Duñabeitia JA, Moret-Tatay C, McGonigal A, Peeters D, Alario F, et al. Smart phone, smart science: how the use of smartphones can revolutionize research in cognitive science. PLoS One 2011;6(9):e24974 [FREE Full text] [doi: 10.1371/journal.pone.0024974] [Medline: 21980370]
7. Timmers C, Maeghs A, Vestjens M, Bonnemayer C, Hamers H, Blokland A. Ambulant cognitive assessment using a smartphone. Appl Neuropsychol Adult 2014;21(2):136-142. [doi: 10.1080/09084282.2013.778261] [Medline: 24826507]
8. Tieges Z, Stíobhairt A, Scott K, Suchorab K, Weir A, Parks S, et al. Development of a smartphone application for the objective detection of attentional deficits in delirium. Int Psychogeriatr 2015 Aug;27(8):1251-1262. [doi: 10.1017/S1041610215000186] [Medline: 25742756]
9. Bajaj JS, Heuman DM, Sterling RK, Sanyal AJ, Siddiqui M, Matherly S, et al. Validation of encephalapp, smartphone-based stroop test, for the diagnosis of covert hepatic encephalopathy. Clin Gastroenterol Hepatol 2015 Oct;13(10):1828-1835.e1 [FREE Full text] [doi: 10.1016/j.cgh.2014.05.011] [Medline: 24846278]
10. Brouillette RM, Foil H, Fontenot S, Correro A, Allen R, Martin CK, et al. Feasibility, reliability, and validity of a smartphone based application for the assessment of cognitive function in the elderly. PLoS One 2013;8(6):e65925 [FREE Full text] [doi: 10.1371/journal.pone.0065925] [Medline: 23776570]
11. Malle C, Ginon B, Bourrilhon C. Brief working memory and physiological monitoring during a high-altitude expedition. High Alt Med Biol 2016 Dec;17(4):359-364. [doi: 10.1089/ham.2016.0022] [Medline: 27548274]
12. Giordano A, Bonometti GP, Vanoglio F, Paneroni M, Bernocchi P, Comini L, et al. Feasibility and cost-effectiveness of a multidisciplinary home-telehealth intervention programme to reduce falls among elderly discharged from hospital: study protocol for a randomized controlled trial. BMC Geriatr 2016 Dec 07;16(1):209 [FREE Full text] [doi: 10.1186/s12877-016-0378-z] [Medline: 27923343]
13. Jelcic N, Agostini M, Meneghello F, Bussè C, Parise S, Galano A, et al. Feasibility and efficacy of cognitive telerehabilitation in early Alzheimer's disease: a pilot study. Clin Interv Aging 2014;9:1605-1611 [FREE Full text] [doi: 10.2147/CIA.S68145] [Medline: 25284993]
14. Green RC, Cupples LA, Go R, Benke KS, Edeki T, Griffith PA, et al. Risk of dementia among white and African American relatives of patients with Alzheimer disease. JAMA 2002 Jan 16;287(3):329-336. [Medline: 11790212]
15. van Osch M, Rövekamp A, Bergman-Agteres SN, Wijsman LW, Ooms SJ, Mooijaart SP, et al. User preferences and usability of iVitality: optimizing an innovative online research platform for home-based health monitoring. Patient Prefer Adherence 2015;9:857-867 [FREE Full text] [doi: 10.2147/PPA.S82510] [Medline: 26170635]
16. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [Medline: 1202204]
17. Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique. Archives de Psychologie 1941;28:286-340.
18. Reitan RM. The relation of the trail making test to organic brain damage. J Consult Psychol 1955 Oct;19(5):393-394. [Medline: 13263471]
19. Stroop JR. Studies of interference in serial verbal reactions. J Exp Soc Psychol 1935 Dec;18(6):643-662. [doi: 10.1037/h0054651]
20. Kirchner WK. Age differences in short-term retention of rapidly changing information. J Exp Psychol 1958 Apr;55(4):352-358. [Medline: 13539317]
21. Zeng EY, Heffner JL, Copeland WK, Mull KE, Bricker JB. Get with the program: adherence to a smartphone app for smoking cessation. Addict Behav 2016 Dec;63:120-124. [doi: 10.1016/j.addbeh.2016.07.007] [Medline: 27454354]

22. Pal R, Mendelson J, Clavier O, Baggott MJ, Coyle J, Galloway GP. Development and testing of a smartphone-based cognitive/neuropsychological evaluation system for substance abusers. J Psychoactive Drugs 2016;48(4):288-294. [doi: 10.1080/02791072.2016.1191093] [Medline: 27260123]
23. van Drongelen A, Boot CR, Hlobil H, Smid T, van der Beek AJ. Process evaluation of a tailored mobile health intervention aiming to reduce fatigue in airline pilots. BMC Public Health 2016 Aug 26;16:894 [FREE Full text] [doi: 10.1186/s12889-016-3572-1] [Medline: 27565140]
24. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and validity of ambulatory cognitive assessments. Assessment 2016 Apr 15 Epub ahead of print. [doi: 10.1177/1073191116643164] [Medline: 27084835]
25. Reitan R. Trail making test. In: Manual for administration and scoring. Tucson: Reitan Neuropsychology Laboratory; 1992.

Abbreviations

AMC: Academic Medical Center
LUMC: Leiden University Medical Center
MMSE: mini-mental state examination
SD: standard deviation
TMT: Trail Making test
15-WVLT: 15-Word Verbal Learning test
CC: correlation coefficient

Edited by G Eysenbach; submitted 04.11.16; peer-reviewed by I Montagni, A Blokland; comments to author 29.01.17; revised version received 26.02.17; accepted 14.03.17; published 25.05.17.

Please cite as:
Jongstra S, Wijsman LW, Cachucho R, Hoevenaar-Blom MP, Mooijaart SP, Richard E
Cognitive Testing in People at Increased Risk of Dementia Using a Smartphone App: The iVitality Proof-of-Principle Study
JMIR Mhealth Uhealth 2017;5(5):e68
URL: http://mhealth.jmir.org/2017/5/e68/
doi:10.2196/mhealth.6939
PMID:28546139

Original Paper

Human-Centered Design Study: Enhancing the Usability of a Mobile Phone App in an Integrated Falls Risk Detection System for Use by Older Adult Users

Richard Harte[1,2], B Eng, Ph.D.; Leo R Quinlan[2,3], BSc, PhD; Liam Glynn[4], MD; Alejandro Rodríguez-Molinero[1,5], MD, PhD; Paul MA Baker[6], PhD; Thomas Scharf[7], FAcSS; Gearóid ÓLaighin[1,2], B Eng, Ph.D.

[1]NUI Galway, Electrical and Electronic Engineering, School of Engineering & Informatics, Galway, Ireland

[2]CÚRAM SFI Centre for Research in Medical Devices, Human Movement Laboratory, NUI Galway, Galway, Ireland

[3]NUI Galway, Physiology, School of Medicine, Galway, Ireland

[4]NUI Galway, General Practice, School of Medicine, Galway, Ireland

[5]Consorci Sanitari del Garraf, Çlinical Research Unit, Vilanova i la Geltrú, Barcelona, Spain

[6]CACP Center for Advanced Communications Policy Georgia Institute of Technology, North Avenue NW, GA 30332, Atlanta, GA, United States

[7]NUI Galway, Irish Centre for Social Gerontology, Institute for Lifecourse and Society, Galway, Ireland

Corresponding Author:
Leo R Quinlan, BSc, PhD
NUI Galway
Physiology
School of Medicine
University Road
Galway,
Ireland
Phone: 353 91 493710
Fax: 353 91 494544
Email: leo.quinlan@nuigalway.ie

Abstract

Background: Design processes such as human-centered design (HCD), which involve the end user throughout the product development and testing process, can be crucial in ensuring that the product meets the needs and capabilities of the user, particularly in terms of safety and user experience. The structured and iterative nature of HCD can often conflict with the necessary rapid product development life-cycles associated with the competitive connected health industry.

Objective: The aim of this study was to apply a structured HCD methodology to the development of a smartphone app that was to be used within a connected health fall risk detection system. Our methodology utilizes so called discount usability engineering techniques to minimize the burden on resources during development and maintain a rapid pace of development. This study will provide prospective designers a detailed description of the application of a HCD methodology.

Methods: A 3-phase methodology was applied. In the first phase, a descriptive “use case” was developed by the system designers and analyzed by both expert stakeholders and end users. The use case described the use of the app and how various actors would interact with it and in what context. A working app prototype and a user manual were then developed based on this feedback and were subjected to a rigorous usability inspection. Further changes were made both to the interface and support documentation. The now advanced prototype was exposed to user testing by end users where further design recommendations were made.

Results: With combined expert and end-user analysis of a comprehensive use case having originally identified 21 problems with the system interface, we have only seen and observed 3 of these problems in user testing, implying that 18 problems were eliminated between phase 1 and 3. Satisfactory ratings were obtained during validation testing by both experts and end users, and final testing by users shows the system requires low mental, physical, and temporal demands according to the NASA Task Load Index (NASA-TLX).

Conclusions: From our observation of older adults’ interactions with smartphone interfaces, there were some recurring themes. Clear and relevant feedback as the user attempts to complete a task is critical. Feedback should include pop-ups, sound tones, color or texture changes, or icon changes to indicate that a function has been completed successfully, such as for the connection

sequence. For text feedback, clear and unambiguous language should be used so as not to create anxiety, particularly when it comes to saving data. Warning tones or symbols, such as caution symbols or shrill tones, should only be used if absolutely necessary. Our HCD methodology, designed and implemented based on the principles of the International Standard Organizaton (ISO) 9241-210 standard, produced a functional app interface within a short production cycle, which is now suitable for use by older adults in long term clinical trials.

(JMIR Mhealth Uhealth 2017;5(5):e71) doi:10.2196/mhealth.7046

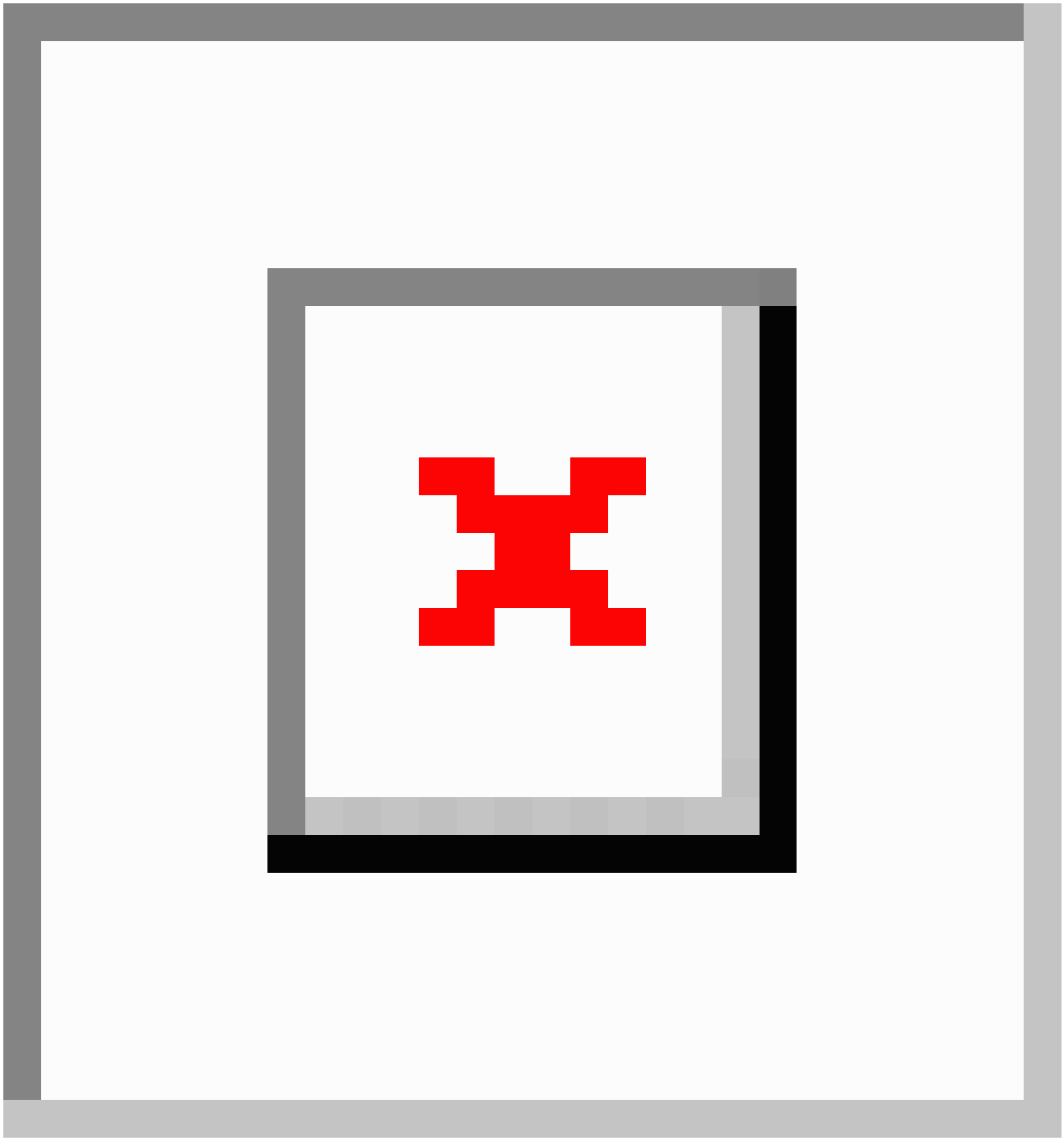
KEYWORDS

human-centered design; user-centered design; human-computer interface; human factors engineering; eHealth; engineering psychology; mHealth

Introduction

Utilizing a human-centered design (HCD) approach, such as that outlined in the International Standards Organization (ISO) 9241-210 [1], during the design of connected health devices ensures that the needs and requirements of the user are taken into consideration throughout the design process. HCD is a multi-stage process that allows for various iterations of a design and subsequent update to the requirements. The importance of involving end users in the design process of health products is recognized, and different approaches have been demonstrated in literature [2-8]. In this paper, we present the implementation of a structured HCD methodology, based on ISO-9241-210, which utilized standard, established techniques to assess and develop the usability and human factors of a smartphone interface with the full involvement of end users and stakeholders. The smartphone interface that was developed and tested is a component of the wireless insole for independent and safe elderly living (WIISEL) system, a system designed to continuously assess fall risk by measuring gait and balance parameters associated with fall risk. The system is also designed to detect falls. The architecture of the system is illustrated in Figure 1. It is proposed that the system can be worn at home by a user for a period of time in order to identify specific gait and balance patterns that may be affecting a user's fall risk. The system is targeted at older adults who represent a high fall risk group. The system consists of a pair of instrumented insoles and a smartphone that are worn by the user. Data collected by embedded sensors in the insoles are sent to the smartphone, where they are then uploaded to a server in a clinic for processing and analysis. The smartphone represents a major interface in the system as this is how the home user will primarily interact with the WIISEL system with the WIISEL app, allowing the user to check the system status, sync with the insoles, send data to their local clinic, and monitor their daily activity.

Figure 1. The wireless insole for independent and safe elderly living (WIISEL) system.



The acquisition and comprehension of information from interfaces can become more difficult as a person progresses into older age. Interfaces in electronic health or medical apps can often be crowded with text and characters, have poor contrast, contain many different colors, and may not present adequate haptic or audio feedback. In terms of visual perception, age-related declines in acuity, contrast sensitivity, and ability to discriminate colors can affect reading rates, character and symbol identification, and button striking accuracy, even with optimal corrections in place [9]. Age-related cognitive decline in domains such as reasoning and memory can affect the ability of the user to comprehend the process they are perceiving on the interface [10]. Deterioration of psychomotor processes such as fine motor control and dexterity can cause problems for users attempting to interact with the physical hardware of the interface [4]. Typically between the ages of 60 and 80 years, individuals can expect up to a 50% decline in visual acuity (particularly in low luminance, low contrast, and glare environments), a reduction in hearing sensitivity by 20dBs, a 14% decline in short-term memory, and a 30% decline in power grip strength, all of which impact how one interacts with computer interfaces [11]. In addition to these physical considerations, older adults can also present a complex user group in terms of attitude toward and previous experience with technology [11].

Methods

A 3-stage HCD methodology was utilized to enhance the usability and user experience of the smartphone app. This methodology was previously described by Harte et al [12].

Phase 1

Use Case Development

The use case document outlined 7 scenarios where the user must directly interact with the smartphone interface. These scenarios were (1) the user logs in to the app, (2) the user syncs the app to the insoles, (3) the user checks the system status, (4) the user uploads the data, (5) the user minimizes the app, (6) the user resets the app, and (7) the user triggers a fall alarm. The use case, which was termed paper prototype version 1, was exposed to 2 groups of stakeholders in the form of structured analysis in order to illicit their feedback [7,13,14].

Expert Use Case Analysis

A total of 10 experts were selected to analyze the use case. The experts were selected from National University of Ireland (NUI), Galway based on their involvement with work related to the use of technology by older adults. We sought multi-disciplinary perspectives, as advised in ISO-92410, and therefore the group consisted of nurses, occupational therapists, physiotherapists, general practitioners, gerontologists, and engineers. The precise expertise of each expert, as well as a self-reported measure of their knowledge of (1) usability and human factors and how it can influence technology use; (2) the end user, their capabilities, and their preferences for technology; and (3) connected health devices that are used in the home can be found in Table 1.

In addition to filling out the Likert statements at the end of each scenario, the expert was instructed to engage in a think-aloud protocol as they walked through each scenario [15]. All feedback was captured by an audio recorder.

End User Representatives Use Case Analysis

A total of 12 older adults were recruited using a typical purposive sample (Inclusion: age 65+ years, community dwelling; Exclusion: profound hearing or vision loss, psychiatric morbidities, and severe neurological impairments) to analyze the use case. The same protocol and interview structure was used to expose the use case document to the older adults and was carried out in the home of the participant. Ethical approval to carry out the interviews and assessments was approved by University Hospital Galway (UHG) research ethics committee. For this analysis, we sought to measure, where applicable, the capabilities a user would call upon to successfully use an interface, so that we could be satisfied that test participants were representative of the target end-user population.

Table 1. Experts involved in use case analysis. Each of the experts was asked to mark out of 10 where they felt their own expertise of usability, the end user, and connected health lay.

#	Profession	Specific experience	End user knowledge	Usability knowledge	Connected health knowledge
1	Clinical researcher in general practice	Industry experience in software design. Research interests include the perception of older adults in the media and the quality of life of dementia sufferers in long stay care.	9	8	7
2	Occupational therapist	Experience in the delivery of occupational health solutions to older adults including ADL[a] assessments, environmental risk assessments, cognitive assessments, and fall prevention strategies.	9	6	3
3	Senior lecturer in nursing	Registered general nurse with a PhD qualification in clinical nursing and has expert experience of treating older adults.	8	8	6
4	GP[b] and senior lecturer	Research addresses chronic disease management and implementing connected health solutions for the management of chronic diseases.	9	5	7
5	GP and head of general practice department	Senior lecturer of general practice and lead researcher in clinical training or teaching practices and methods, as well as workplace learning and development.	9	6	4
6	Psychology researcher	Holds a PhD in psychology with research interest in team situation awareness in critical environments and designing instructional materials. Currently working in the area of examining lifestyle and technology factors associated with gestational diabetes mellitus.	7	8	7
7	Clinical researcher in general practice	Former practising nurse currently a masters researcher pursuing projects in connected health and telehealth solutions in rural communities.	8	6	8
8	GP and senior lecturer in general practice	HRB[c] Cochrane Fellow currently practicing as a GP with expert experience of treating older adult patients. Research interests are in multimorbidity with a focus on connected health solutions.	10	6	8
9	IT[d] lecturer and expert in user-centered design	IT researcher specialising in human computer interaction. Research interests heavily focused on the employment of user-centered design techniques for mobile devices.	6	8	4
10	Geriatrician and professor of geron-technology	MD specializing in geriatrics and PhD qualification in preventive medicine and public health. Has expert experience of treating older adults as well as specific research interests in epidemiology, geron-technology, and tele-health care.	10	8	8
		Average expert group knowledge of key areas.	8.4	7	6.4

[a]ADL: activities of daily living.

[b]GP: general practitioner.

[c]HRB: health research board.

[d]IT: information technology.

We measured the cognitive and visual capabilities of the user and the components of the processes we measured are illustrated in Figure 2.

We used a short battery of standardized tests to measure each of the capabilities presented in Figure 2. The tests and their relevance to the analysis are listed in Table 2.

High contrast acuity (HCA) was measured using a Snellen chart at a distance of 3m. Low contrast acuity (LCA) was measured for 5% and 25% contrast using SLOAN letter charts at a distance of 3m. Standardized illumination was provided for these 2 tests using a light box from Precision Vision (precision-vision.com). Constrast sensitivity (CS) was measured using a MARS chart at a distance of 40cm, whereas low contrast acuity in low luminance (LCALL) was measured with a SKI chart at a distance of 40cm. Color discrimination (CD) was measured using a Farnsboro D-15 test. Reading acuity (RA) was measured using a Jaeger chart at a distance of 40cm. Each participant also completed 2 cognitive performance tests based on the Whitehall study [22]. Spatial reasoning was assessed using the Alice Heim 4-I (AH4-I). The AH4-I tests inductive reasoning, measuring one's ability to identify patterns, and to infer principles and rules [24]. Short-term memory was assessed with a 20-word free recall test. Expected values of each test per age group and the actual measured can be found in Tables 3 and 4.

Figure 2. Physiological capabilities required to interact with use case.

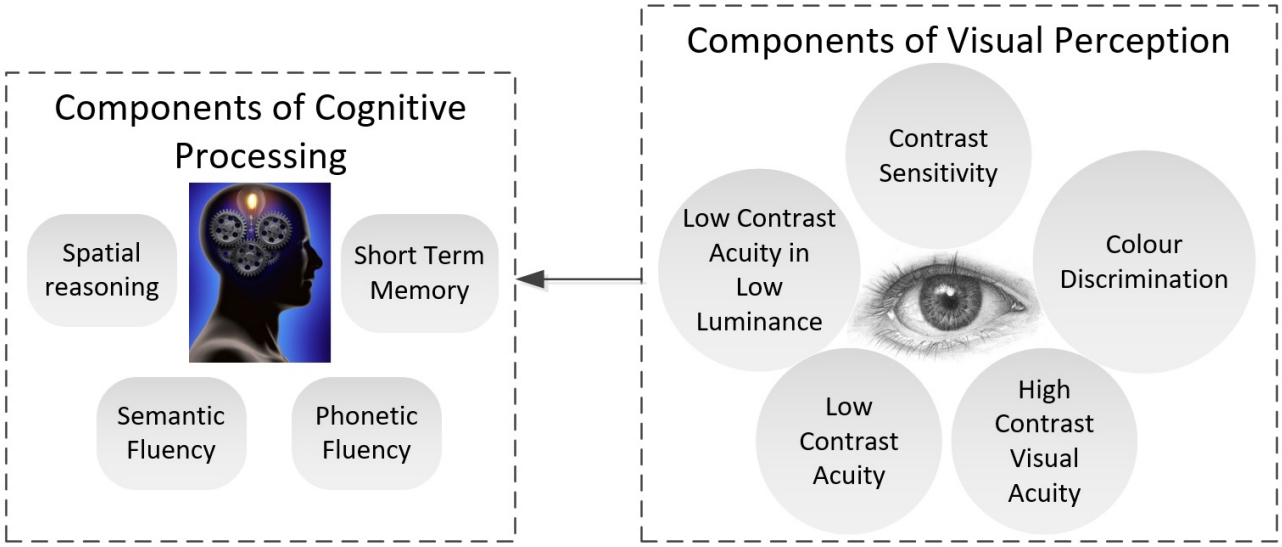


Table 2. Battery of tests.

Interactive process	Measure	Meaning and relevance
Visual perception		
	High contrast acuity	A general measure of visual capability and the ability to discern spaces between characters on a 100% contrast interface [16].
	Reading acuity	A measure of acuity when reading full words on an interface [17].
	Low contrast acuity	A general measure of visual capability and the ability to discern spaces between characters on a 5% and 25% contrast interface [18].
	Contrast sensitivity threshold	The contrast threshold at which the user can successfully identify a character [19].
	Color discrimination	Ability to discriminate colors on an interface [20].
	Low contrast acuity in low luminance	The ability to discern spaces between characters on a low contrast and poorly illuminated surface.
Cognitive processing		
	Spatial reasoning	The ability to interpret space on an interface and infer relationships between elements has been cited as a major component of website usability and software interfaces in general [9,21,22].
	Short-term memory	Memory, specifically short-term memory has been cited as an important factor in one's ability to maintain visual attention of an interface [22,23].

Table 3. Average visual performance metrics measured and split by age group. The average is compared with the expected score for that age group. Data presented in each column as expected or measured

Age (years)	n (number of participants who fall into age group)	HCA[a] (expected or measured)	RA[b] (expected or measured)	LCA[c] (expected or measured)	LCALL[d] (expected or measured)	CS[e] (expected or measured)	CD[f]
61-65	1	1/1	1/1	0.67/0.79	0.33/0.41	1.68/1.8	No defects
66-70	2	1/1	1/1	0.62/0.71	0.29/0.27	1.54/1.55	No defects
71-75	3	0.91/0.83	0.91/0.8	0.49/0.64	0.25/0.22	1.42/1.33	1 participant with very mild blue yellow confusion (tritanopia)
76-80	5	0.83/0.88	Not applicable	0.4/0.54	0.2/0.22	1.2/1.42	1 participant with very moderate blue yellow confusion (tritanopia)
81-85	1	0.76/0.66	Not applicable	0.3/0.5	0.17/0.2	0.61/0.7	No defects

[a]HCA: high contrast acuity.

[b]RA: reading acuity.

[c]LCA: low contrast acuity.

[d]LCALL: low contrast acuity in low luminance.

[e]CS: color sensitivity.

[f]CD: color discrimination.

Table 4. Expected scores and mean measured scores for cognitive tests for all 12 participants. The average is compared with the expected score for that age group. Data presented in each column.

Age (years)	n (number of participants who fall into age group)	Spatial reasoning (range 0-65) (expected or measured)	Short-term memory (range 0-20) (expected or measured)
61-65	1	(30-46)/30	(6.21-6.43)/8
66-70	2	(30-46)/38.5	(4.79-5.74)/6.5
71-75	3	(29.7-40)/35	(4.7-5.5)/6.3
76-80	5	(29.7-40)/29.8	(4.3-5.4)/5.9
81-85	1	(25-40)/26	(4-5.1)/4

Identification and Categorisation of Usability Problems

The audio feedback acquired during the analysis of the use case document by the experts and end users was “intelligently” transcribed [25] and clearly defined usability problems were extracted from the transcript. All of the problems identified by each expert, and end user were collated for each scenario. All problems were documented and illustrated in a structured usability and human factors problems report [26] and were accompanied by selected testimony from a corresponding expert or end user who elaborated on the nature of the problem for the purpose of the design team. This report was analyzed by system designers who provided potential solutions to each problem where possible.

Phase 2

In response to the feedback from phase 1, a new paper prototype was developed (paper prototype version 2) and made available for expert inspection. A working version of the app with accompanying user manuals was also developed on a Google Nexus 5 smartphone (working prototype version 1) and made available for expert walkthrough. We returned to the original experts and carried out a 2-part usability inspection. First, the experts inspected the solutions to the problems they had identified in phase 1 using a new version of the use case (paper prototype version 2) as a guide. This use case only presented the problems that the experts identified in their original analysis and showed how the problems had been addressed. Second, they inspected the prototype app (working prototype version 1) utilizing a cognitive and contextual walkthrough methodology.

Phase 3

The new manuals and updated interface (working prototype version 2) were exposed to the 10 older adults who had previously analyzed the use case (2 of the 12 subjects who had originally analyzed the use case were unavailable in phase 3 testing). After measuring the time taken to complete each task and the number of errors made, the after scenario questionnaire (ASQ) and the NASA Task Load Index (NASA-TLX) were administered to the participant after the task was completed. The ASQ is a Likert scale that interrogates a user’s perception of efficiency, ease of use, and satisfaction with manual support [27]. The NASA-TLX is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on 6 subscales: (1) mental demands, (2) physical demands, (3) temporal demands, (4) own performance, (5) effort, and (6) frustration [28].

Results

This section presents the summary of results from each phase, as well as the changes made to the interface and support documentation after each phase.

Phase 1: Use Case Analysis (Paper Prototype Version 1)

The combined expert analysis and end user analysis identified 21 problems. We have provided 13 examples of problems, which are presented in Table 5. These 13 problems were chosen for illustration because they represent unique problems, the other 8 problems were considered repetitions or derivatives of the other 13, and therefore, we felt it was not important to describe them. The problem ID number assigned to each problem was used for the remainder of the design process to allow for easier problem tracking throughout the process.

The problems from Table 5 are presented in Table 6 in order of severity rating based on the mean Likert scores assigned by the experts. The maximum individual score that was given by the 10 experts is also included to highlight the fact that some experts may have given a more severe rating than what the mean or standard deviation indicates. The heuristic category to which each problem belongs is also included.

Table 5. List of identified problems and which use case scenario it was identified in.

Problem ID number	Problem description (use case scenario)
1	The difference in operation between the home button and back button is not clear (user minimizes app)
2	Overall login sequence (user must log in to the app)
3	Buttons on keypad are too small for this population (user must log in to the app)
4	WIISEL[a]icon not prominent enough on app menu (user must check the system status)
5	Having to upload the data will be too hard to remember to do (uploading data by exiting app)
6	Feedback during the process is not clear or may cause anxiety (uploading data by exiting app)
7	No prompt to indicate to the user that a manual connection is now required (user must connect to the insoles)
8	Colors are too similar in places (uploading data by exiting app)
9	Feedback regarding connection status is unclear (user connects to insoles using app)
10	Homescreen information is not clear (user must check the system status)
11	Options presented are not clear (fall alarm or notification)
12	App text is too small (user must check the system status)
13	Buttons on exit screen need to be bigger (uploading data by exiting app)

[a]WIISEL: wireless insole for independent and safe elderly living.

Table 6. Problems uncovered by experts and rated based on mean Likert scores.

Problem ID number	Heuristic category	Severity rating (0-4) (σ)	Max severity rating given (0-4)
1	Cognitive directness	2.5 (1.2)	4
2	Consistency and compliance of task structure	2.4 (1.1)	4
3	Discernibility (button size)	2.2 (1.3)	4
4	Discernibility (icons)	2.2 (1.3)	4
5	Consistency and compliance of task structure	2.1 (0.9)	3
6	Completeness and sufficiency of meaning	2.1 (1)	4
7	Consistency and compliance of task structure	1.9 (0.6)	4
8	Discernibility (color tone and contrast)	1.9 (1.2)	4
9	Completeness and sufficiency of meaning	1.7 (0.9)	4
10	Completeness and sufficiency of meaning	1.5 (0.8)	4
11	Consistency and compliance of task structure	1.4 (1)	3
12	Discernibility (text size)	1.3 (0.75)	3
13	Button size (discernibility)	1.2 (0.9)	4

Table 7. Problems uncovered by end users and rated based on mean Likert scores.

Problem ID number	Heuristic category	Severity rating (0-4) (σ)	Max severity rating given (0-4)
1	Cognitive directness	1.83 (0.89)	3
2	Consistency and compliance of task structure	1.5 (0.7)	2
4	Discernibility (button size)	1.5 (0.8)	2
7	Discernibility (text size)	1.33 (1)	3
13	Discernibility (button size)	1.2 (0.9)	3
8	Discernibility (color tone and contrast)	1.15 (0.6)	2
9	Completeness and sufficiency of meaning	1(1.2)	3
6	Completeness and sufficiency of meaning	0.91 (0.6)	3
12	Discernibility (text size)	0.91 (0.7)	3

The older adult end user analysis found 14 problems, all of which were problems that had been identified by the expert group (the same problem ID number is used). Of the 13 problems listed in Table 6, 9 were uncovered by end users. These are presented in Table 7 in order of severity (as in Table 6).

Testimony from experts and users alike were used to provide insight into the problems and help designers better understand the problem. Themes were sought from the transcripts to uncover which characteristics of the interface experts and users most commonly found problematic. For example, regarding the login sequence for the smartphone app:

> *If not absolutely necessary this sequence should be removed from the use of the phone. At the very least it should be made sure that this only needs to be carried out by the clinician in the clinic once.*
>
> *Maybe a voice password could be used or simply a pin number that only requires numerical values and does not require an email address.*

Insufficient screen feedback and prompts for the user when carrying out certain tasks was identified as a recurring theme:

> *There should be a prompt to upload the data. When he (the user) presses the back button it should prompt the user that the data is about to be uploaded. The warning sign on the Exit pop-up box will cause anxiety and should be avoided.*
>
> *I suggest that the interface should have one indicator saying if everything is working OK and if not, the interface should say specifically what the issue is.*
>
> *The battery icon needs to change colour/shape when it is decreasing.; There needs to be a message which appears on the screen telling the user to initiate this (connection) sequence (PLEASE PRESS HERE TO ATTEMPT CONNECTION) and an indicator on the screen should tell them where to press.* [Recommended by Expert 8]

The size of screen elements such as icons, buttons and text were identified as being problematic:

> *(Made in reference to the pop up boxes in particular, for example "Invalid mail or Password" during login,) the screen needs to be utilised better, pop up boxes need to be bigger and more prominent.*
>
> *There is no reason why the large screen space could not be utilised more effectively for these buttons (referring to exit pop-up buttons).* [Expert 1]
>
> *This (referring to an icon in top left hand corner to show that the app is running) is a good idea, but it is just too small for older adult users.*

The results of the expert analysis and the end user analysis were compiled separately and then were presented in a problem report for system developers, with all problems listed with severity ratings and related testimony. The developers returned a proposal on how each problem could be solved, which were then reviewed by the usability engineering team. Examples of proposals that were accepted by the usability engineers are shown in Table 8.

Not all identified problems could be easily fixed by the system developers. Some aspects of the interface were built into the Android operating system (OS) and therefore could not be changed, whereas some problems could not be solved within the time constraints of the project. Where it was clear that the developer could not affectively address a problem through interface changes, the usability team proposed an alternative as to how the problem severity could be at least reduced if not completely eliminated. Some of examples of these problems are shown in Table 9.

Table 8. Problems that were directly addressed by system developers.

Problem ID number	System developer comments
2	The login will be a once off action carried out at the clinic to simply match the data coming through to the patient who is using the app. We have debugged the app so that any crashes should not mean the user has to log back into the app (login cookie is stored on phone cache). We will also make it so that the user can see the password as they are typing to decrease the chance of error, as suggested by the experts.
4	We will change this to a more prominent symbol that will be slightly bigger although is constrained by the operating system. We will make this symbol the same icon as the app icon.
6	We will change the feedback text to “Are you sure you want to close this application? After closing, the data will be sent to the server.” We will also change the caution symbol to an Information symbol based on your suggestion.
8	Contrast has been increased and text size increased to make it more prominent against the dark background.
9	We will remove the text “connect in 10 seconds pop-up” and just have “auto connection started” and “an everything is ok” pop-up once sequence is complete.
10	The “timer” text has been removed. We will also introduce colors for the symbols, red when the symbol is not in the ideal state, and green when it is.
11	We will introduce a green and red button choice with related symbols.
12	Text size will be increased and some redundant components will be removed from the interface to make more space.

Table 9. Problems that could not be directly addressed by system developers and which in turn had a proposed solution by the usability team.

Problem ID number	Problem	Scenario	System developer comments	Usability team proposal
1	The difference in operation between the home button and back button is not clear	User minimizes app	This is an Android design and cannot be changed and we feel that adding another button (an exit button) to the interface may cause further confusion	We will provide an instruction sheet that will show the user clearly the difference between the 2 buttons, emphasizing in particular that the back button is only used for uploading the data
3	Buttons on keypad are too small for this population	User logins to the app	This is an Android design and cannot be changed. The only solution would be to “buy” another keypad design that will be expensive	Short tutorials will be conducted for users on how to effectively use the keypad at the onset of use to improve confidence
5	Having to upload the data will be too hard to remember to do	Uploading data by exiting app	At this stage of development, an automatic data push is not feasible but will be considered for future	We will emphasize this scenario in our user manuals to reflect the fact that it needs to be carried out periodically
7	No prompt to indicate to the user that a manual connection is now required	User must connect to the insoles	We will improve the auto connection and introduce an option in the settings to turn off auto connection	We will describe the sequence in the short form manual, with steps for when a user should attempt a manual connection

Update of Paper Prototype and Development of First Working Prototype

Based on this communication between the development team and the usability engineers, a working app prototype for the Google Nexus 5 smartphone was developed as well as a full set of user manuals based on the use cases and the feedback from the use case analyses. The use case was also updated to reflect the changes to the interfaces. Figures 3 and 4 show examples of how the updated interface (paper prototype version 2) compares with the paper prototype version 1. In Figure 3, we see how color indicators have been introduced to enhance the feedback on the system status screen. Text size has been increased and some elements have been removed from the interface to reduce crowding. Figure 4 shows how the login screen has been updated with a decrypted password as well as increased text size and button size.

Figure 3. (a) The old interface showing the system status. Experts did not like the dull colors and crowded interface. Some users did not like the fact that there was no change of colors to indicate low battery, weak signal etc; (b) The updated interface with color indicators for connection, signal strength, and battery life, as well as increased text size and contrast.

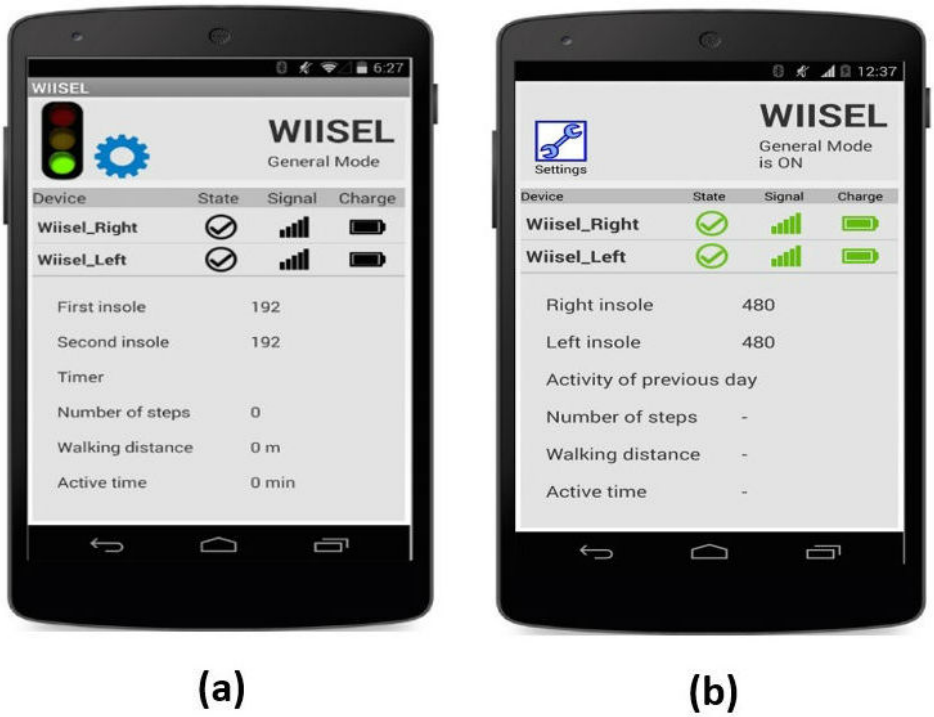
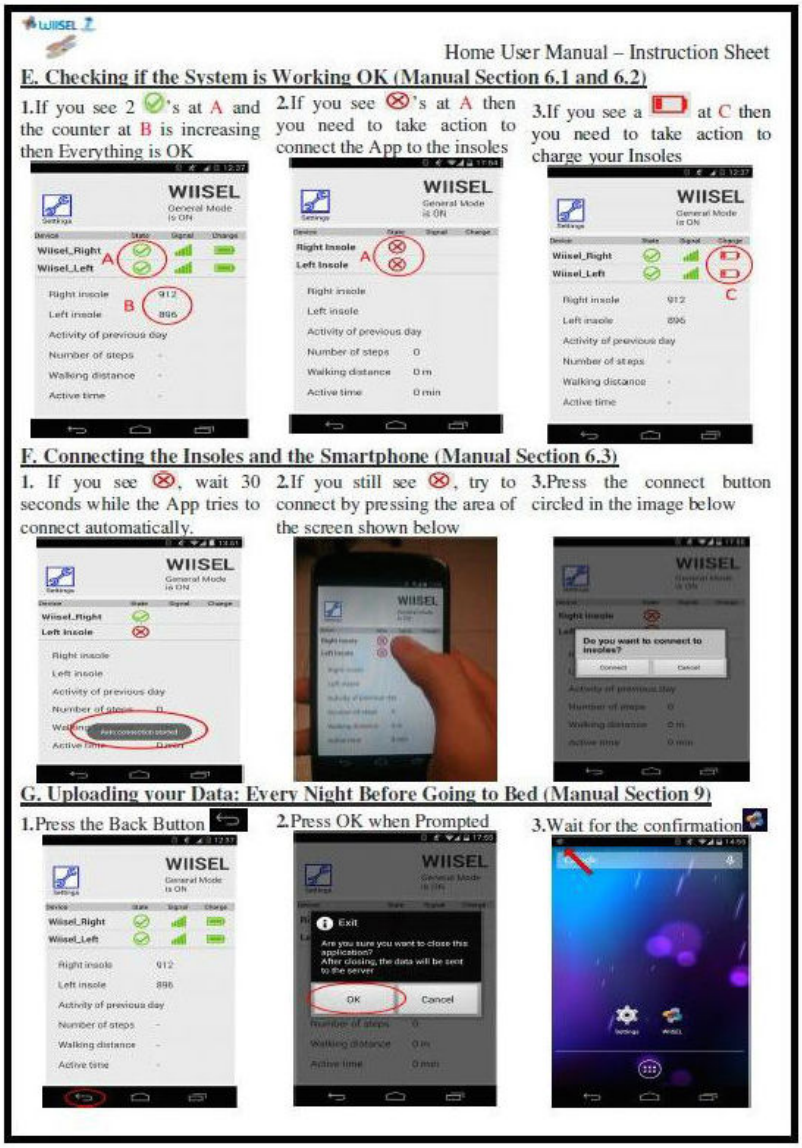


Figure 4. (a) Experts were concerned with the small button size and the fact that the password was encrypted meaning an older adult might lose their place when typing. This problem was also identified by end users; (b) Increased text size and a larger, more prominent sign in button as well as a decrypted password.

Figure 5. One side of the basic instruction sheet (short form manual) describing the connection and uploading sequences.



Where the problems identified by the experts could not be addressed by an interface change, user manuals were created to offset any confusion of difficulties the user might encounter with the interface. In order to create an effective user manual, the original use case was updated with all the interface changes made by designers. Each use case scenario now became a section of the user manual with the same chronological order maintained where applicable. For example, the use case scenario where the user connects to the insoles became a "how to connect" section in the user manual and was followed by a "how to upload" section, as in the use case. Two forms of manual were created, a short form manual entitled the "basic instruction sheet" which contained basic instructions on a double-sided laminated sheet, and a longer form manual laid out in similar style to the use case that elaborated on the instructions provided in the basic instruction sheet and provided additional instructions for procedures that would not be considered routine. Another version of these 2 forms were also created for clinicians with additional information on how to set up the system for the user, change settings, calibrate insoles, and adjust fall detection settings. A selected sections of the manual is presented in Figure 5.

Phase 2: Expert Inspection Results

Use Case Inspection of Paper Prototype Version 2

Table 10 presents examples of how the various problems uncovered during the use case analysis in phase 1 were addressed and compares the problem rating it received from the first use case analysis (paper prototype version 1) with the new rating it received from the analysis of the updated interface in phase 2 (paper prototype version 2).

The inspection found that of the 21 original problems identified by the experts, 3 had now received a rating of 0 from the experts, 17 had received decreased ratings, and 1 (ID# 11) had received an increased rating.

Table 10. Comparison of problem ratings between paper prototype V1 problems and the updated interface (paper prototype V2). The max individual score that was given by the 10 experts is also included to highlight the fact that some experts may have given a more severe rating than the mean or standard deviation indicates.

Problem ID number	Severity rating (0-4) (σ) Paper prototype version 1	Max severity rating given (0-4)	How was the problem addressed?	Severity rating (0-4) (σ) Paper prototype version 2	Max severity rating given (0-4)
1	2.5	4	A manual section was added that explained the operation of each button in the context of overall phone operation and in the context of the WIISEL[a]app.	1.4	2
2	2.4	4	Debugging of the app and improved connection sequence means that app resets are not as likely, leading to a decreased need for the user to have to login. Button size was increased and the password decryption during the sequence was removed.	0.3	1
5	2.1	3	Additional manual information was added and instructions on setting a daily reminder on the phone.	1.4	2
6	2.1	4	The caution symbol has been replaced with an information symbol, additional text information has been added explaining to the user what is happening regarding the data upload.	0.4	1
11	1.5	3	Red and green have been introduced as “I have fallen” option (red) and “I am Ok” option (green). Whereas experts agree with the notion of illustrations and color coding, they are now concerned that there is no text labels on the buttons. One expert pointed out that red could be confused for a cancel button (ie, to cancel the alarm) in the same way as it would be when answering a phone call. This could lead to a user accidentally sending a fall alert to carer during a false positive sequence in which the user is forced to press a button in a hurry.	2.1	4
10	1.5	4	The addition of the green indicators for “good” and orange and red indicators for “bad” such as for the battery symbol have been welcomed.	0.1	1
12	1.3	3	Whereas the homescreen interface had improved, some experts felt that some space was not being utilized well and that small text and crowding was still an issue.	0.44	1

[a]WIISEL: wireless insole for independent and safe elderly living.

Table 11. Average metrics and consensus for 9 experts. After scenario questionnaire (ASQ) scores range from 1-7, where 1 is the most satisfied and 7 is the least satisfied the user can be.

Scenario	Time taken (s)	Average errors made	ASQ[a] score	Comments
Check the system status	4.7	0.5	2	The increasing numbers (referring to the incrementing counters) on the interface are still unclear to some experts. Whereas the experts acknowledge that this indicates "data is streaming," the indication should be that it is either connected or it is not, any other information than that is completely redundant. Documentation is a little crowded, would like to see more space given in the manual
Connect to the insoles	48.0	0	3	This task but could be made easier by giving more feedback to the user and simplifying the interface somewhat. If the connection takes a couple of minutes then the user needs to be made aware that something is happening or else they will just keep pressing the connect button, possibly causing a crash or accidentally disconnecting it. The ambiguities in the connection sequence need to be made clear in the manuals, ''don't panic, give the system a chance etc.''
Upload data	4.3	0	2	There is a concern that there is no immediate feedback to let the user know they have completed the task successfully. The manual indicates that an icon will appear in the top left hand corner of the screen, however, it is obvious that this does not appear straight away if there is a lot of data, this should be made clear in the manual or just removed, as it may cause anxiety.
User minimizes app			4	The difference in operation between the back button and the home button, while addressed, is not made completely clear in the user manual. This will be important for users particularly if they intend to user other functions on the phone.
Reset app	23.4	0.7	3	Not a very straight-forward sequence given the number of screens that need to be navigated, but under supervision this should be OK. This is quick if the user knows what they are looking for, although they could get easily lost. The user manual should explain to the user that they made need to scroll down in each menu to reach the option they need. If the user does not see the exact same screen that they see in the user manual they will think something is wrong.
Login to app	27.0	1.1	3	This will present challenges, particularly the keyboard. If the user can follow the manual then it will be easy but any digression from the main path will cause problems. The time is OK, although mistakes with the user credentials will obviously increase the time as well as the user frustration. More steps need to be added to this sequence in the manual.
Respond to fall alarm	7	0.3	3	This is an easy sequence but the confusion over the options makes it a little bit more burdensome especially on users with any form of cognitive impairment. Very quick to do, provided the user is clear on what option they are pressing. The documentation here is inadequate and needs to explain the situations in which each option may need to be pressed.

[a]ASQ: after scenario questionnaire.

Expert Cognitive or Contextual Walkthrough With Working Prototype Version 1

Table 11 shows the captured average metrics from each scenario, with the time and errors made metric captured. Accompanying the metrics are a selection of comments from experts.

Of the 8 scenarios, three achieved a score of "satisfied"and four achieved a score of "somewhat satisfied," whereas one achieved a neutral score. No scenarios scored a perfect score of 1, indicating that all scenarios require some improvement, particularly regarding the clarity and flow of the supporting documentation. These data are best illustrated in a radar chart (Figure 6). A radar chart allows for multiple data series to be displayed across common variables, each variable having its own axis (the dotted line). The axis values go from low to high as you read toward the center of the chart, with lower scores indicating a better outcome (data points near the edge of the chart). The chart in Figure 6 shows how the 3 individual components of the ASQ score, satisfaction with ease of completion, time taken, and effect of supporting documentation.

In response to comments by the experts during the inspection, the user manuals were updated, and several minor changes were made to the interface. These updates are listed in Table 12.

Figure 6. All basic scenarios scored consistently well regarding ease of completion (blue) with just slight superficial changes, the more challenging scenarios such as login and reset registered higher (worse) scores. Only one scenario, connection routine, scored poorly in the time taken (red) metric, owing to the length of time it takes the insoles to sync with the app. Several experts were confused by some of the layout and instructions in the manuals (green), with improvement required for several scenarios, particularly the instructions for the fall alarm sequence.

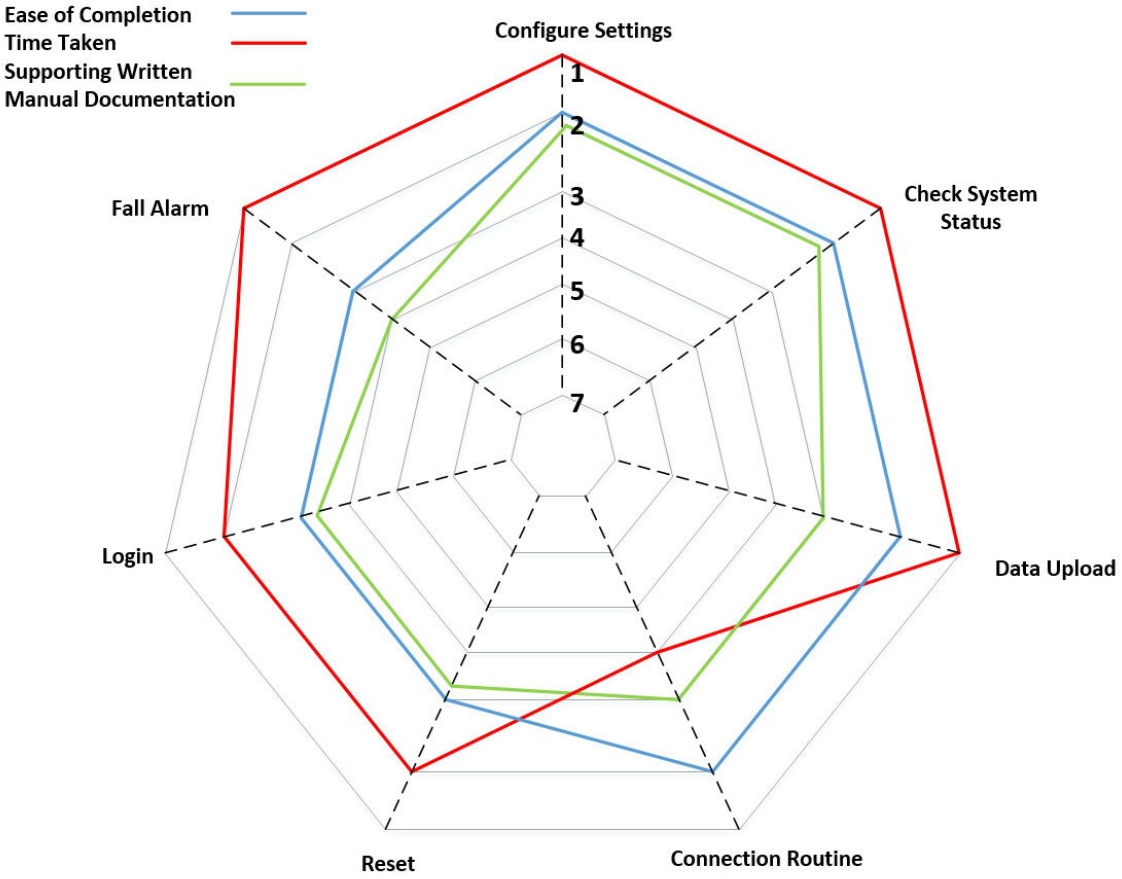


Table 12. Changes made to the user manuals and interface based on expert inspection.

Scenario	Suggestions	Changes made
Check the system status	This section of the manual needs to be less crowded	The documentation now includes 6 steps instead of the original 4. A step is included to explain that the app may take several seconds to start up and how to lock the phone again. Increased text size
Connect to the insoles	Would like to see some explanation of the crash sequence in the user manual.	The same number of steps is maintained with additional labels indicating where on the screen the user may have to press during the connection sequence. Increased button or text size
Upload data	Manual indicates that an icon will appear in the top left hand corner of the screen, however, it is obvious that this does not appear straight away if there is a lot of data, this should be made clear in the manual or just removed, as it may cause anxiety.	The third step, which explained that an icon would appear upon successful completion has been removed to avoid confusion as it does not always appear straight away. The section now also includes further explanation of what the back button is used for. Increased button or text size
User minimizes app (home button)	The difference in operation between the back button and the home button, while addressed, is not made completely clear in the user manual.	A section explaining the function of this button was placed on the same page as the section explaining the use of the back button. This was done in order to provide a clear distinction between the function of the 2 buttons
Reset app	The user manual should explain to the user that they made need to scroll down in each menu to reach the option they need. If the user does not see the exact same screen that they see in the user manual, they will think something is wrong.	Expanded from a 3-step instructional process to a 5-step process. A section was also introduced to explain to the user how to best interact with the touchscreen in terms of scrolling and striking
Login to app	More steps need to be added to this sequence in the manual	Expanded from a 4-step process to a 6-step process including additional instructions on how to access the number keypad and find the @ symbol Increased button size
Respond to fall alarm	The documentation here is inadequate and needs to explain the situations in which each option may need to be pressed. As regards the options, it is suggested that red and green not be used to distinguish options and that text labels also be used for the buttons to accompany and supplement images	Expanded from 1-step to a 3-step process with clear illustrations to show when the user might experience New buttons introduced to indicate cancel and confirm Increased button size

Figure 7. (a) Fall alarm interface before expert inspection, the red and green caused confusion as the red was associated with "cancel" as you would find on a phone call interface; (b) Fall alarm interface after expert inspection, a more appropriate symbol was introduced for the help button whereas the cancel button was changed to a more neutral blue with appropriate labeling.

These changes led to working prototype version 2 and a new set of user manuals that now contained 4 laminated sheets. Figure 7 shows an example of how the fall alarm interface has been updated.

Phase 3: Usability Testing With End Users

Table 13 shows the average metrics for the 10 test participants during the usability testing of working prototype version 2, whereas Figure 8 illustrates the breakdown of the ASQ metric in terms of satisfaction with ease of completion, time taken, and support documentation (there was some confusion with the reset and login sequences in the user manual (green) which is explained further in Table 13).

The results of the NASA-TLX was performed on paper and the metrics are shown in Table 14. A score of 100 indicates maximum burden on the user, whereas a score of 0 indicates no burden. The first 4 tasks scored very well, indicating little to no burden on the user. The login and reset procedures, due to the number of steps involved, created the most mental, physical, and effort burden, as well as the most frustration, particularly the login procedure. The most temporal burden was created by the fall alarm procedure, due to the timer on the screen, forcing the user to make a hasty choice.

Table 13. Performance metrics for each scenario during user testing with working prototype 2, with related commentary as observed during the testing. The after scenario questionnaire (ASQ) score ranges from 1-7, where 1=best score possible and 7=worst score possible.

Scenario	Time taken	Errors made	ASQ[a]	Comments
Check the system status	19	0.4	1	All users found this very easy to complete and manuals clear to follow. The only errors encountered were when users released the screen slide lock too early, which occurred with 4 of the 10 users.
Connect to the insoles	31	0.13	1	Whereas users found the procedure and manual easy to follow, the time taken for the sync to complete caused minor frustration. The only error encountered were when some users held the manual connection button for too long.
Upload data	13	0.13	1	All users found this very easy to complete and manuals clear to follow. Some minor errors included pressing the cancel button instead of the OK button. Whereas the OK button was clearly marked as the button to press in the user manual, sometimes the user would press cancel without consulting the manual.
Reset app	112	1.0	2	While quite a complex sequence, most user's found it easy to complete, but were susceptible to minor errors, such as accidentally pressing the wrong menu option, or accidentally pressing while scrolling. These errors are down to unfamiliarity with touch screen interfaces and "heavy handedness." There was one error with regards to the layout of the manual.
Login to app	171	0.88	2	This sequence took the most time, due to most user's unfamiliarity with touchscreen keypads. There was a huge disparity in times, ranging from 30s to nearly 5 min, with those who had previous experience with smartphones faring generally better. The manual layout also caused some confusion with user's having to jump a step to find out how to enter numbers and then having to return to the previous step.
Respond to fall alarm	6	0.5	2	The original fall sequence caused an error for every second user, who thought the red option was the cancel option, as you would expect on a mobile phone call.
Respond to fall alarm	6	0	1	The new alternative fall sequence proved more successful, with the removal of the red or green option causing less confusion with no errors reported.

Table 14. NASA Task Load Index (NASA-TLX) scale breakdown by scenario. The NASA-TLX score ranges from 1-100, where 1=worst score possible and 100=best score possible.

Scenario	Overall score	Mental	Physical	Temporal	Performance	Effort	Frustration
Check the system status	4.9	3.8	8.5	4.3	4.3	4.7	4.0
Connect to the insoles	7.0	9.9	5.6	12.9	4.4	4.4	4.9
Upload data	4.1	3.2	4.0	3.3	5.3	3.3	5.2
Minimize app	3.6	4.7	3.7	2.8	3.0	3.8	3.5
Reset app	30.3	42.3	16.3	15.0	17.5	53.3	37.2
Login to app	39.1	54.7	29.0	20.3	20.7	65.8	43.8
Respond to fall alarm 1	26.6	43.5	7.7	59.5	20.2	18.3	10.5
Respond to fall alarm 2	13.6	22.8	6.2	33.3	4.8	6.5	7.7
Most burdensome scenario	Login	Login	Login	Respond to fall alarm 1	Login	Login	Login

Figure 8. All scenarios scored maximum for ease of completion (blue) apart from the fall alarm 1 which caused slight confusion. Time taken (red) was not considered a major issue for any of the scenarios, with the connection routine not scoring maximum due to the nature of the syncing process, whereas the unfamiliarity with typing caused some users to mark down the login sequence. There was some confusion with the reset and login sequences in the user manual.

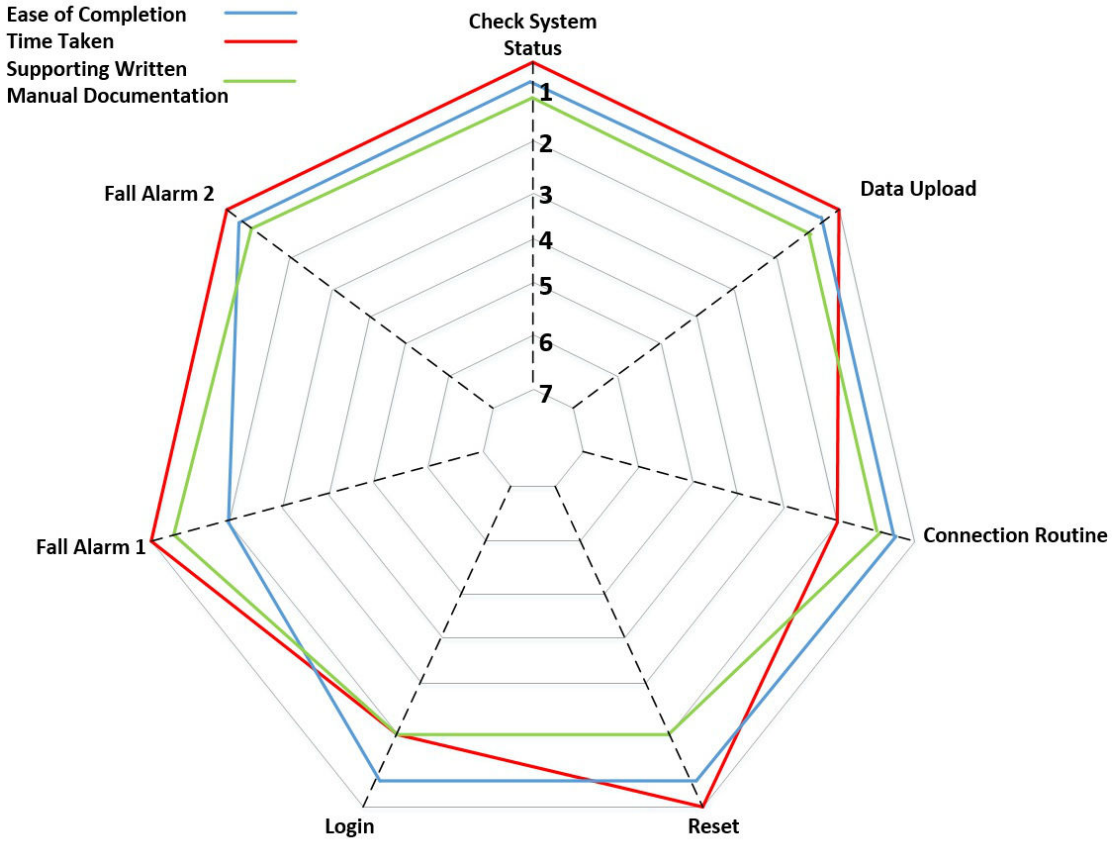


Table 15. Likert items severity rating (range 0-4, 0=no problem, 4=most severe problem) for interface ergonomics by scenario. Some Likert items did not apply to certain scenarios. An x indicates that there was no Likert statement for that particular interface aspect for that scenario.

Scenario	Color	Text	Buttons	Keypad buttons	Icons size	Icon meaning
Check the system status	0	0.25	x	x	0.12	0
Connect to the insoles	x	x	0.12	x	x	x
Upload data	0	0.12	0.12	x	x	x
Login to app	0	0.12	0	0.37	x	x
Respond to fall alarm 1	0	0	0	x	x	x
Respond to fall alarm 2	0	0	0	x	x	x

Table 16. Presents the evolution of three distinct problems through the testing lifecycle with the usability metrics taken at each stage.

Problem ID number	Phase 1	Phase 2		Phase 3		
	Severity rating (0-4) (σ)	New severity rating after inspection	Expert ASQ[a] score (0-7)	End user ASQ score (0-7)	NASA-TLX (0-100)	Caused error during user testing?
2	1.77	0.6	3	2	39	On average, users made 0.88 errors during this scenario
6	1.41	0.4	3	1	4	On average, users made 0.13 errors during this scenario
11	1.09	1.2	3	1	13	On average users made 0 errors during this scenario

[a]ASQ: after scenario questionnaire.

Table 17. System usability scale (SUS) metric, split into overall usability and learnability, captured at each phase.

	Phase 1		Phase 2		Phase 3	
	Learnability $\bar{x}$ (σ)	SUS[a] total $\bar{x}$ (σ)	Learnability $\bar{x}$ (σ)	SUS total $\bar{x}$ (σ)	Learnability $\bar{x}$ (σ)	SUS total $\bar{x}$ (σ)
Experts	35 (24.23)	55 (23.6)	48.75 (26.36)	68.75 (11.6)	n/a	n/a
End users	58 (26.43)	78 (10.77)	n/a	n/a	87.5 (5)	88 (3.75)

[a]SUS: system usability scale.

Table 15 shows the Likert response for different aspects of the interface in each scenario. The severity rating is calculated in the same manner as phase 1 and 2.

Summary of Results

With combined expert and end-user analysis of a comprehensive use case having originally identified 21 problems with the system interface, we have only seen observed 3 of these problems in user testing (problem ID 1, 2, and 12). Satisfactory ASQ ratings were obtained during validation testing by both experts and end users, and final testing by users shows the system requires low mental, physical, and temporal demands according to the NASA-TLX. Table 16 shows how three of the problems (problems involving flow, consistency, and feedback) have evolved over the testing cycle. Problem 2 and 6 show a clear linear improvement from phase 1-3, with problem 2 an example of a problem that despite best efforts remained a cause of potential user frustration due to the unfamiliar style of touchscreen keyboards. Problem 6 represents an example of a problem that was effectively mitigated through interface changes and manual support. Problem 11 is an example of a problem that was actually exasperated by an interface change, causing greater confusion to users, although this was effectively identified and mitigated between phase 2 and 3.

The system usability scale (SUS) metrics after each phase are presented in Table 17. The SUS is split into 2 scales: (1) overall usability and (2) learnability [29]. Early phases showed widely variable SUS scores, particularly among experts, whereas phase 3 scores showed agreement among end users that the interface had achieved some level of acceptability.

Discussion

Overview

We have presented a multi-phase, mixed-method HCD approach to improve the user experience of a smartphone interface, which forms part of a connected health system. Our approach was designed to uncover and mitigate any usability problems as early as possible, before they were exposed to end users during usability testing and in formal clinical trials. This paper presents one full cycle of our HCD process, with each phase representing an iteration where a design update or refinement took place. Our approach has met the specific recommendations for a HCD process [30]. We have adopted the input of multi-disciplinary skills and perspectives by eliciting the feedback of both an end-user group and an appropriately experienced expert group throughout the process. We have sought to gain an explicit understanding of users, tasks, and environments and consideration of the whole user experience through the adoption of a use case that provided context of use for system tasks and scenarios and through the examination of the perceptual and cognitive needs of the target end user. We utilized a user-centered evaluation driven design using standard usability evaluation metrics at each point in the cycle. We involved users throughout the design process, at both early and later stages. Finally, we employed an iterative process, split into 3 stages or phases that allowed for user feedback to be worked into design updates.

Principal Findings

From our observation of older adults' interactions with smartphone interfaces, there were some recurring themes. Clear and relevant feedback as the user attempts to complete a task is critical (in line with contemporary literature) [31,32]. Feedback should include pop-ups, sound tones, color or texture changes or icon changes to indicate that a function has been completed successfully, such as for the connection sequence (problem ID# 9). For text feedback, clear and unambiguous language should be used so as not to create anxiety, particularly when it comes to saving data such as in the data upload sequence (problem ID# 6). Older adults not familiar with technology are often afraid that they might delete something by accident or fail to save important data properly. Warning tones or symbols, such as a caution symbol, should only be used if absolutely necessary. For audio feedback, clear and low frequency tones should be used. Login sequences where the user is required to input text with a QWERTY keyboard should be avoided (problem ID 2), particularly for those who have no previous touchscreen experience. If a login sequence is considered necessary for security or identification purposes, it should be ensured that a login process is made as simple as possible (do not hide password, be clear about what username is required, supply ample support documentation for process). For simple interface elements, text sizes should be at least 10pts (Didot system), whereas button sizes should have a surface area of no less than approximately 200mm^2 [11,33].

In terms of metrics, we used 4 different subjective measurement systems (Likert scales, ASQ, NASA-TLX, and SUS) to assess the usability of the interface at different stages. The Likert scales allowed for quick satisfaction ratings of the perceived ease of use of each task in the use case and of the suitability of interface elements such as text and button size. The ASQ was more suitable for postscenario ratings when the user had actually completed the task, whereas the NASA-TLX was used to supplement the ASQ to provide further details on what kind of burden, be it physical or cognitive, the task placed on the user. The SUS was utilized when the user had completed a full use of the system and carried out all tasks. We observed that all of

these metrics are providing the similar information, just in slightly different resolutions, and that a mixture of metrics allows us different insights into user perceptions of usability. For example, in phase 3, from looking at the ASQ scores of the login sequence, we could conclude that the user was satisfied with the ease of the task. However, when we looked at the NASA-TLX scores, we observed that the task was creating a large mental demand on them. These 2 metrics, whereas showing us seemingly conflicting pieces of information, may be telling us that the user judged the task as being easy simply because they completed it successfully, regardless of the difficulty they encountered or the time it had taken them. It is only when they think about the task in terms of the NASA metrics that they become honest about what kind of burden the task placed on them. The SUS was a useful general indicator of overall usability but its wide variability (Table 17) suggests that it is best used with larger sample sizes. High SUS scores do not guarantee the system will not suffer usability problems in the field [34]. These metrics are probably best used to supplement more objective metrics such as task times and error rates.

Procedural Observations

In terms of efficiency, our methodology proved to be successful. The utilization of the use case analysis activities during phase 1 provided a focus for all stakeholders on the context of and the intended use of the system. The time it took for each individual to analyze and provide feedback was on average 1 h. Within this hour, the individual was experiencing and commenting on context, was being formally interviewed, was filling out questionnaires, and was providing opinions on interface concepts. Therefore in one session the use case analysis provides multiple streams of data, whereas in previous literature, this kind of feedback would need to be gathered across multiple activities, such as surveys, interviews, and ethnographic observations. In phase 2, the use of expert inspection groups also proved highly efficient. We recommend that research groups and design teams maintain an inspection group who can carry out on hand inspections of new system versions. This group, which can comprise 4-6 members, need not necessarily be qualified usability engineers but can be trained in techniques such as heuristic evaluations and cognitive walkthroughs. In terms of how long it took to complete each phase, as this was a case study as part of a research project, the amount of time spent on each phase was probably drawn out longer than it would be in a more industrial setting. In all, the 3 phases together took approximately 12 months, with phase 1 taking the bulk of the time (approximately 6 months) as use cases were developed and redeveloped and end users were interviewed and tested. After the app was developed and testable, the phases became shorter, with phase 2 and 3 taking approximately 3-4 months each. As the methodology is applied in future, it will become more refined, allowing for quicker development cycles.

Limitations

Time and technology constraints meant that not all design requirements could be implemented. For example, the replacement of the manual data upload with an automatic periodic data upload could not be implemented in time by the engineering team. Similarly, the structure of the Android OS meant that some user and expert recommendations could not be implemented, particularly regarding the positioning of pop-ups or the nature of data storage. Some design changes led to a decrease in user experience, particularly for the fall alarm sequence (problem ID# 11). It became clear during user testing that the use of red and green in an emergency situation may not be the best practice, with some users confusing the red emergency button for a cancel button, like it may be presented on a phone call screen (red for “hang-up”). In this case, the design team failed to take into account the recommendation of one expert who predicted that a red or green option may cause confusion. We can conclude from this that taking on board opinions from different stakeholders can present a challenge for designers. However, the nature of our iterative methodology meant that this problem was identified and addressed between phase 2 and 3.

In phase 1, the older adult end users tended to be very optimistic about how they would handle the system and the smartphone interface, overall giving higher scores in response to Likert statements and for the overall SUS score. Experts tended to be more pessimistic but this was probably due to their vast experience with older adults and technology. Most experts conceded that the use case analysis was a hypothetical one and that the capabilities of the older adult population are extremely variable, however, they felt that it was an extremely useful exercise in identifying major potential problems and addressing them early in the design process. Despite the difference in outlook between the experts and older adults, both groups reached agreement on most problems, particularly about the perceived difficulty of the login process and the lack of clear feedback when checking the system status and during the data upload process. We can conclude from this that utilizing multiple perspectives from different groups is an important feature of a good human-centered design process.

Conclusions

The HCD Methodology we have designed and implemented based on the principles of ISO 9241-210 has produced a functional app interface that is now suitable for exposure to older adults in long term clinical trials. We have applied appropriate testing techniques given the context of the interface being assessed. We would consider this a thorough and robust method for testing and informing design changes of all types of interactive connected health systems.

Acknowledgments

This work was part funded by the EU FP7 project Wireless Insole for Independent and Safe Elderly Living (WIISEL), project number FP7-ICT-2011-288878.

Authors' Contributions

The methodology for this study was conceived and designed by RH, LRQ, and GOL. The experiments were carried out by RH with the support of LG, TS, and ARM, each of whom contributed both usability and medical knowledge to the testing. The data was compiled and analyzed by RH, LRQ, and GOL and reviewed by LG, ARM, and PMAB. All authors contributed equally to the introduction and discussion sections of the paper. The paper as a whole was reviewed and edited where necessary by all authors before submission.

Conflicts of Interest

None declared.

References

1. International Organization for Standardization. ISO. 2010. Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems URL: https://www.iso.org/standard/52075.html [accessed 2017-05-18] [WebCite Cache ID 6qXpjolRx]
2. Martínez-Pérez B, de la Torre-Díez I, Candelas-Plasencia S, López-Coronado M. Development and evaluation of tools for measuring the quality of experience (QoE) in mHealth applications. J Med Syst 2013 Oct;37(5):9976. [doi: 10.1007/s10916-013-9976-x] [Medline: 24014255]
3. Shah Syed Ghulam Sarwar, Robinson I, AlShawi S. Developing medical device technologies from users' perspectives: a theoretical framework for involving users in the development process. Int J Technol Assess Health Care 2009 Oct;25(4):514-521. [doi: 10.1017/S0266462309990328] [Medline: 19845981]
4. Sesto ME, Irwin CB, Chen KB, Chourasia AO, Wiegmann DA. Effect of touch screen button size and spacing on touch characteristics of users with and without disabilities. Hum Factors 2012 Jun;54(3):425-436. [Medline: 22768644]
5. Abugabah AJ, Alfarraj O. Issues to consider in designing health care information systems: a user-centred design approach. E-Journal of Health Inform 2015;9(1):8 [FREE Full text]
6. Borycki E, Kushniruk A, Nohr C, Takeda H, Kuwata S, Carvalho C, et al. Usability methods for ensuring health information technology safety: evidence-based approaches. contribution of the IMIA working group health informatics for patient safety. Yearb Med Inform 2013;8:20-27. [Medline: 23974544]
7. Vermeulen J, Neyens JC, Spreeuwenberg MD, van RE, Sipers W, Habets H, et al. User-centered development and testing of a monitoring system that provides feedback regarding physical functioning to elderly people. Patient Prefer Adherence 2013 Aug;7:843-854 [FREE Full text] [doi: 10.2147/PPA.S45897] [Medline: 24039407]
8. Developing an insole for elderly fall prevention. WIISEL: Wireless Insole for Independent and Safe Elderly Living URL: http://www.wiisel.eu/ [accessed 2017-03-16] [WebCite Cache ID 6p0O6YP86]
9. Echt KV, Burridge AB. Predictors of reported internet use in older adults with high and low health literacy: the role of socio-demographics and visual and cognitive function. Phys Occup Ther Geriatr 2011;29(1):23-43.
10. Wagner N, Hassanein K, Head M. The impact of age on website usability. Comput Hum Behav 2014;37:270-282.
11. Harte RP, Glynn LG, Broderick BJ, Rodriguez-Molinero A, Baker PM, McGuiness B, et al. Human centred design considerations for connected health devices for the older adult. J Pers Med 2014 Jun 04;4(2):245-281 [FREE Full text] [doi: 10.3390/jpm4020245] [Medline: 25563225]
12. Harte R, Glynn L, Rodriquez-Molinero A, Baker PM, Scharf T, Quinlan LR, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017;4(1):e8. [doi: 10.2196/humanfactors.5443] [Medline: PMC5374275]
13. Hull E, Jackson K, Dick J. Requirements Engineering. London: Springer; 2011.
14. Khajouei R, Peute LW, Hasman A, Jaspers MW. Classification and prioritization of usability problems using an augmented classification scheme. J Biomed Inform 2011 Dec;44(6):948-957 [FREE Full text] [doi: 10.1016/j.jbi.2011.07.002] [Medline: 21782036]
15. Cooke L. Assessing concurrent think-aloud protocol as a usability test method: a technical communication approach. IEEE Trans Profess Commun 2010 Sep;53(3):202-205 [FREE Full text] [doi: 10.1109/TPC.2010.2052859]
16. Balcer LJ, Galetta SL, Polman CH, Eggenberger E, Calabresi PA, Zhang A, et al. Low-contrast acuity measures visual improvement in phase 3 trial of natalizumab in relapsing MS. J Neurol Sci 2012 Jul 15;318(1-2):119-124. [doi: 10.1016/j.jns.2012.03.009] [Medline: 22521274]
17. Findl O, Leydolt C. Meta-analysis of accommodating intraocular lenses. J Cataract Refract Surg 2007 Mar;33(3):522-527. [doi: 10.1016/j.jcrs.2006.11.020] [Medline: 17321405]
18. Pineles SL, Birch EE, Talman LS, Sackel DJ, Frohman EM, Calabresi PA, et al. One eye or two: a comparison of binocular and monocular low-contrast acuity testing in multiple sclerosis. Am J Ophthalmol 2011 Jul;152(1):133-140 [FREE Full text] [doi: 10.1016/j.ajo.2011.01.023] [Medline: 21570055]
19. Thayaparan K, Crossland MD, Rubin GS. Clinical assessment of two new contrast sensitivity charts. Br J Ophthalmol 2007 Jun;91(6):749-752 [FREE Full text] [doi: 10.1136/bjo.2006.109280] [Medline: 17166891]

20. Schneck ME, Haegerstrom-Portnoy G, Lott LA, Brabyn JA. Comparison of panel D-15 tests in a large older population. Optom Vis Sci 2014 Mar;91(3):284-290 [FREE Full text] [doi: 10.1097/OPX.0000000000000152] [Medline: 24535417]
21. Sofi F, Valecchi D, Bacci D, Abbate R, Gensini GF, Casini A, et al. Physical activity and risk of cognitive decline: a meta-analysis of prospective studies. J Intern Med 2011 Jan;269(1):107-117 [FREE Full text] [doi: 10.1111/j.1365-2796.2010.02281.x] [Medline: 20831630]
22. Singh-Manoux A, Kivimaki M, Glymour MM, Elbaz A, Berr C, Ebmeier KP, et al. Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. BMJ 2012;344:d7622 [FREE Full text] [Medline: 22223828]
23. Bundesen C, Habekost T, Kyllingsbæk S. A neural theory of visual attention and short-term memory (NTVA). Neuropsychologia 2011 May;49(6):1446-1457. [doi: 10.1016/j.neuropsychologia.2010.12.006] [Medline: 21146554]
24. Heim AW. SJDM. 1968. AH4 group test of intelligence URL: http://www.sjdm.org/dmidi/AH4_intelligence_test.html [accessed 2017-03-16] [WebCite Cache ID 6p0QH0QLn]
25. Isaac C. Weloty. 2017. Intelligent verbatim transcription URL: https://weloty.com/intelligent-verbatim-transcription/ [accessed 2017-03-16] [WebCite Cache ID 6p0PnHSwZ]
26. Sears A, Jacko JA. Human-Computer Interaction: Development Process. New York: CRC Press; 2009.
27. Lewis JR. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. ACM SIGCHI Bulletin 1991;23(1):78-81.
28. Hart SG. Development of NASA-TLX (Task Load Index): results of empirical theoretical research. Adv Psychol 1988;52:139-183. [doi: 10.1016/S0166-4115(08)62386-9]
29. Lewis JR, Sauro J. The factor structure of the system usability scale. 2009 Presented at: Proceedings of the 1st International Conference on Human Centered Design; 2009; Berlin, Heidelberg.
30. Giacomin J. What is human centred design? Design J 2014 Dec 01;17(4):606-623. [doi: 10.2752/175630614X14056185480186]
31. Fisk D. Designing for older adults. In: Principles and Creative Human Factors Approaches. New York: CRC Press; 2009.
32. Zhou J, Rau PP, Salvendy G. Use and design of handheld computers for older adults: a review and appraisal. Int J Hum Comput Interact 2012 Dec;28(12):799-826. [doi: 10.1080/10447318.2012.668129]
33. Jin ZX, Plocher T, Kiff L. Touch screen user interfaces for older adults: button size spacing. : Springer; 2007 Presented at: Universal Access in Human Computer Interaction. Coping with Diversity; 2007; Berlin p. 933-941. [doi: 10.1007/978-3-540-73279-2_104]
34. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008;24(6):594.

Abbreviations

ADL: acitivities of daily living
ASQ: after scenario questionnaire
CD: color discrimination
CS: color sensitivity
GP: general practitioner
HCA: high contrast acuity.
HCD: human-centered design
HCI: human-computer interation
HRB: health research board
IT: information technology
LCA: low contrast acuity
LCALL: low contrast acuity in low luminance
OS: operating system
RA: reading acuity
SUS: system usability scale
UHG: University Hospital Galway

Edited by A Keepanasseril; submitted 24.11.16; peer-reviewed by E Afari-kumah, R Berenbaum; comments to author 27.12.16; revised version received 21.03.17; accepted 21.03.17; published 30.05.17.

Please cite as:
Harte R, Quinlan LR, Glynn L, Rodríguez-Molinero A, Baker PMA, Scharf T, ÓLaighin G
Human-Centered Design Study: Enhancing the Usability of a Mobile Phone App in an Integrated Falls Risk Detection System for Use by Older Adult Users
JMIR Mhealth Uhealth 2017;5(5):e71
URL: http://mhealth.jmir.org/2017/5/e71/
doi:10.2196/mhealth.7046
PMID:28559227

Original Paper

Incorporation of a Stress Reducing Mobile App in the Care of Patients With Type 2 Diabetes: A Prospective Study

Maya Munster-Segev[1], RD, BSc, MPH; Oren Fuerst[2], PhD; Steven A Kaplan[3], MD; Avivit Cahn[1,4], MD

[1]Diabetes Unit, Department of Internal Medicine, Hadassah Hebrew University Hospital, Jerusalem, Israel

[2]Eco-Fusion, Herzliya, Israel

[3]Icahn School of Medicine and Director of the Men's Wellness Program Mount Sinai, New York, NY, United States

[4]Endocrinology and Metabolism Unit, Department of Internal Medicine, Hadassah Hebrew University Hospital, Jerusalem, Israel

Corresponding Author:
Avivit Cahn, MD
Diabetes Unit
Department of Internal Medicine
Hadassah Hebrew University Hospital
PO Box 12000
Jerusalem, 91120
Israel
Phone: 972 26778021
Fax: 972 26425097
Email: avivit@hadassah.org.il

Abstract

Background: Severe and sustained emotional stress creates a physiological burden through increased sympathetic activity and higher energy demand. This may lead to increased oxidative stress and development of the metabolic syndrome. Emotional stress has been shown to contribute to the onset, progression, and control of type 2 diabetes (T2D). Stress management and biofeedback assisted relaxation have been shown to improve glycemic control. Use of a mobile app for stress management may enhance the scalability of such an approach.

Objective: The aim of this study was to assess the effect of using a mobile app of biofeedback-assisted relaxation on weight, blood pressure (BP), and glycemic measures of patients with T2D.

Methods: Adult patients with T2D and inadequate glycemic control (hemoglobin A1c [HbA1c]>7.5%) were recruited from the outpatient diabetes clinic. Baseline weight, BP, HbA1c, fasting plasma glucose (FPG), triglycerides (TG), and 7-point self-monitoring of blood glucose were measured. Patients were provided with a stress reducing biofeedback mobile app and instructed to use it 3 times a day. The mobile app—Serenita—is an interactive relaxation app based on acquiring a photoplethysmography signal from the mobile phone's camera lens, where the user places his finger. The app collects information regarding the user's blood flow, heart rate, and heart rate variability and provides real-time feedback and individualized breathing instructions in order to modulate the stress level. All clinical and biochemical measures were repeated at 8 and 16 weeks of the study. The primary outcome was changes in measures at 8 weeks.

Results: Seven patients completed 8 weeks of the study and 4 completed 16 weeks. At week 8, weight dropped by an average of 4.0 Kg (SD 4.3), systolic BP by 8.6 mmHg (SD 18.6), HbA1c by 1.3% (SD 1.6), FPG by 4.3 mmol/l (4.2), and serum TG were unchanged.

Conclusions: Stress reduction using a mobile app based on biofeedback may improve glycemic control, weight, and BP.

(JMIR Mhealth Uhealth 2017;5(5):e75) doi:10.2196/mhealth.7408

KEYWORDS

diabetes mellitus, type 2; biofeedback; physiological stress response; mobile health; telemedicine

Introduction

Type 2 diabetes (T2D) is a chronic disease which has reached epidemic proportions worldwide. Recent data show a total of 415 million patients in the world suffering from diabetes with an expected rise to 642 million by 2040 [1]. Factors contributing to this alarming rise in the prevalence of diabetes include obesity compounded by a sedentary lifestyle, increasing life-expectancy, urbanization, reduced physical activity, increased sugar consumption, and low fruit and vegetable intake [1].

Severe and sustained emotional stress creates a physiological burden through increased sympathetic activity and higher energy demand. This may lead to increased oxidative stress and development of conditions such as the metabolic syndrome, accelerated aging, and cardiovascular disease [2]. Emotional stress has been shown to be a significant accelerator to the development and progression of chronic diseases in general, and T2D in particular [3-5]. Furthermore, individuals experiencing increased levels of stress may find it more difficult to manage their chronic condition, often leading to its exacerbation and resulting in increased stress.

Mind-body interventions facilitate autonomic flexibility, enhance self-regulation, and induce a "relaxation response" which is characterized by parasympathetic dominance, reduced sympathetic activation, and increased heart rate variability. Yoga includes a wide range of mind-body practices including postures, breathing, meditation, and relaxation practices that are reported to counteract physiological stress. Compared with nonyoga practitioners, regular yoga practitioners are reported to have lower heart rate, breath rate, blood pressure (BP), metabolic rate, and higher heart rate variability. Regular yoga practice inspires a sense of psychological and physiological equilibrium that is different from rest, physical relaxation, and sleep [2]. Yoga has been found to be beneficial in reducing oxidative stress in T2D [6].

Incorporation of stress management [7,8] as well as biofeedback-assisted relaxation [9] in the care of diabetes to facilitate increased patient compliance and improved glycemic control has been previously described.

Serenita is a mobile app that engages the body's natural relaxation response by using a personalized breathing exercise. In this study, we assessed the effect of incorporating this novel stress management mobile app in the care of patients with diabetes.

Methods

Study Oversight

This was a single center study conducted in in the diabetes unit of Hadassah Hebrew University Medical Center. The study was approved by the Institutional Ethics Board. All subjects provided signed informed consent (NCT 02691273).

Study Population

The study included patients aged 18 years or above with T2D duration of at least one year who were willing to comply with the study protocol. Patients' hemoglobin A1c (HbA1c) was >7.5%, and they had not attained their glycemic target as assessed by the investigators. Access to a mobile phone—either Android version 4 and up or iPhone 4 and up—with iOS 7 and up was mandatory.

Patients with type 1 diabetes, pregnant women, or those suffering from an acute disease or a chronic significantly decompensated disorder were excluded from the trial.

Study Outcomes

This was a single-arm pilot study designed to assess the impact of using a stress reducing mobile app on the glycemic control, weight, and BP of patients with T2D. The main outcomes were changes in HbA1c, triglycerides (TG), fasting plasma glucose (FPG), 7-point daily glycemic profile, weight, and blood from baseline to week 8. Compliance with the use of the mobile app was to be assessed as well. Patients additionally completed a quality of life questionnaire (SF-12) [10], as well as questionnaires regarding multiple aspects of daily life including sleeping hours, mood, exercise, and hunger at baseline and at week 8. Measures were repeated at week 16.

Study Conduct

Patients attending the outpatient diabetes clinic and willing to participate were screened for the trial. Eligible patients were enrolled in the study.

At baseline, medical history was obtained and a physical exam was conducted. Vital signs, weight, BP, and blood test were taken. The patients were requested to complete a 7-point glucose profile in the 2 days before baseline visit. Dietary counseling in accordance with the American Diabetes Association (ADA) clinical nutrition guidelines was delivered by a registered dietician [5]. Each patient received detailed instructions on how to install and operate the mobile app. Patients were requested to fill out the SF-12 and the questionnaire regarding daily life routine.

The patients received subsequent weekly phone calls to assess their compliance with the mobile app, dietary recommendations, and any further impending issues for 8 weeks, and then monthly calls till week 16. At 8 and 16 weeks after baseline, patients were invited to the clinic where anthropometric measures, blood tests, and 7-point glucose measurements were reassessed, in addition to the reinforcement of dietary guidelines and completion of the questionnaires.

Description of the Mobile App

The mobile app—Serenita—is an interactive relaxation app based on acquiring a photoplethysmography signal from the mobile phone's camera lens, where the user places his finger [11]. Information related to the user's blood flow, heart rate, and heart rate variability is extracted. These signals are then filtered and processed using a large array of algorithms and machine learning to assess the user's physiological stress level and subsequently recommend real-time stress reducing interactive breathing patterns.

Table 1. Baseline and follow-up weight and blood pressure.

Patient #	Visit	Weight (kg)	Change from baseline	Systolic BP[a] (mmHg)	Change from baseline	Diastolic BP (mmHg)	Change from baseline
1	Baseline	79		156		80	
	Week 8	76.6	−2.4				
	Week 16	77.5	−1.5	148	−8	81	1
3	Baseline	110.2		126		84	
	Week 8	103.2	−7				
	Week 16	100	−10.2	111	−15	69	−15
4	Baseline	105.6		140		83	
	Week 8	102.1	−3.5	151	11	84	1
	Week 16	101.4	−4.2				
8	Baseline	122.2		148		86	
	Week 8	110	−12.2	118	−30	75	−11
	Week 16	109.5	−12.7				
11	Baseline	74.2		137		90	
	Week 8	74.4	0.2	122	−15	79	−11
	Week 16						
12	Baseline	89.1		133		74	
	Week 8	86.7	−2.4	113	−20	71	−3
	Week 16						
14	Baseline	62		121		88	
	Week 8	61.5	−0.5	131	10	84	−4
	Week 16						
Average (SD)[b] (week 8)	N	7		5		5	
	Baseline	91.8 (21.7)		135.6 (9.8)		84.2 (6.4)	
	Week 8	87.8 (17.9)	−4.0 (4.3)	127.0 (15.2)	−8.6 (18.6)	78.7 (5.8)	−5.5 (5.1)
Average (SD) (week 16)	N	4		2		2	
	Baseline	104.3 (18.2)		141.0 (21.2)		82.2 (3.1)	
	Week 16	97.1 (13.7)	−7.2 (5.2)	129.3 (26.4)	−11.7 (5.2)	75.0 (8.5)	−7.2 (11.5)

[a]BP: blood pressure.

[b]SD: standard deviation.

Serenita encourages deep diaphragmatic breathing, often referred to as yogic breathing, as well as coherence breathing—breathing synchronized to the heart rate. Additionally, the app offers the user real-time feedback on the effect of their breathing pattern on additional physiological parameters such as heart rate and heart rate variability and accordingly reflects to the individual their current stress level. The mobile app optimizes the pattern of breathing so that the user can assess his or her physiological stress level and obtain a personalized breathing pattern to rapidly reduce the level of stress and increase focus. Figure 1 shows screen caps from the app.

Patients were instructed to use the mobile app 3 times a day for 10 min each session.

Statistical Analyses

This study was designed as a pilot study, and sample size determination was not planned to meet any specific significance and power requirements. Descriptive analyses were conducted for all measurements in the study.

Figure 1. Screen cap from the mobile app.

Results

Fifteen patients were screened for the study and 9 patients were enrolled. Causes for screen failure included low HbA1c (3), incompatible mobile phone (1), lack of home WiFi connection (1), and personal reasons (1). Of the 9 patients enrolled, 7 completed the study. One was unable to operate the app, as the heart rate signal was not adequately transmitted (possibly due to thick skin in his fingers), and one was lost to follow-up after the baseline visit. Seven patients attended the visit at week 8 and 4 attended the visit at week 16. The average age was 55 (SD 11.6), 71% (5/7) were male, diabetes duration was 11 years (SD 7.3), and 71% (5/7) had been using insulin. Baseline, visit 8, and visit 16 weight, BP, HbA1c, FPG, and TG are shown in Tables 1 and 2. At week 8, weight dropped by an average of 4.0

Kg (SD 4.3) from a baseline of 91.8 Kg (SD 21.7) (n=7); systolic BP dropped in average by 8.6 mmHg (SD 18.6) (n=5); HbA1c was reduced by an average of 1.3% (SD 1.6) from 9.0% (SD 0.7) at baseline; FPG was reduced by 4.3 mmol/l (SD 4.2) (mean 77.4 mg/dl, SD 75.6), and serum TG were unchanged (n=6). Baseline medications remained unchanged throughout the trial, except for a 30% increase in insulin dose of patient #11.

Figure 2. Self-monitoring blood glucose levels of the individual patients. Legend—self-monitoring of blood glucose (SMBG) levels. Solid lines—baseline, dashed line—week 8 visit, dotted line—week 16 visit. Patient 14 did not complete SMBG.

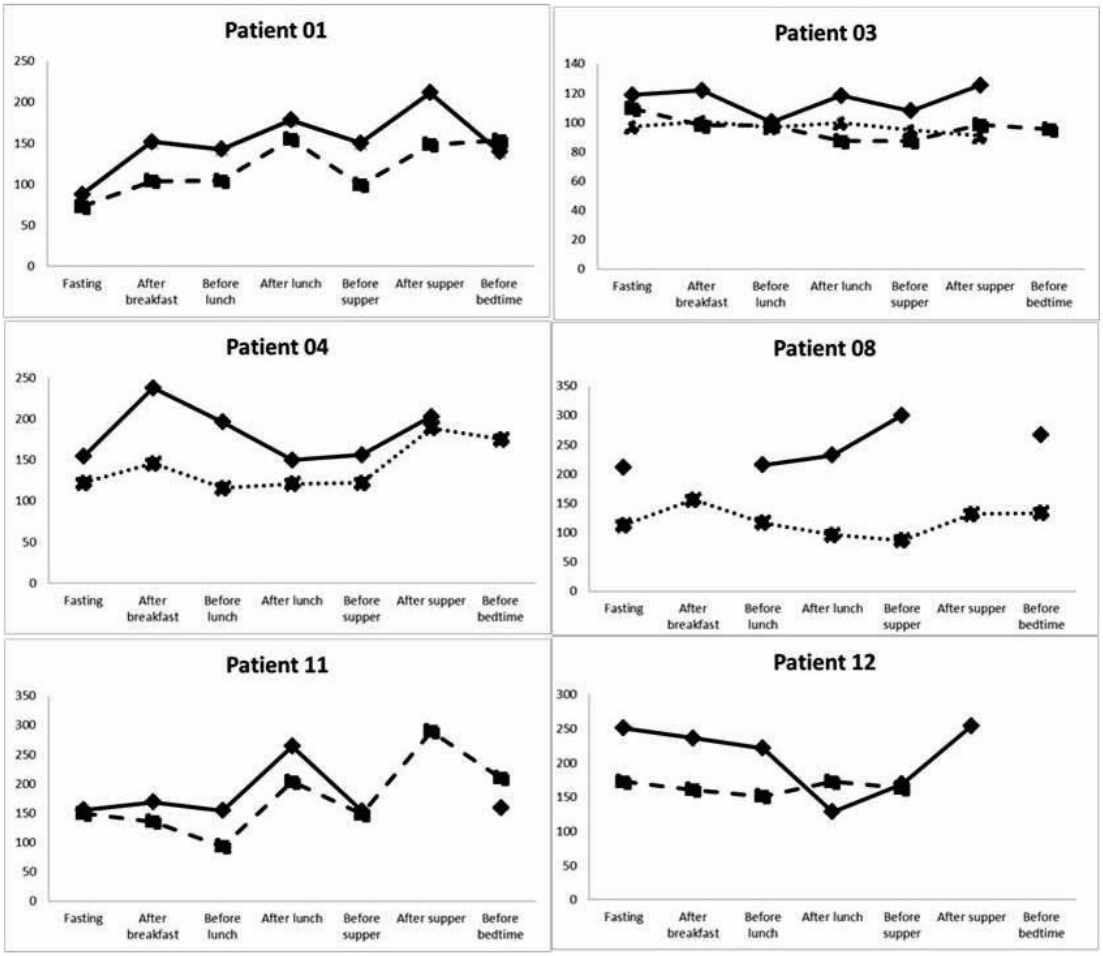


Table 2. Baseline and follow-up biochemistry analyses.

Patient #	Visit	HbA1c[a] (%)	Change from baseline	FPG[b] (mmol/l)	Change from baseline	TG[c] (mmol/l)	Change from baseline
1	Baseline	7.6		8.1		1.5	
	Week 8	6.3	−1.3	5.5	−2.6	1.6	0.1
	Week 16	6.2	−1.4	3.5	−4.6	1.8	0.3
3	Baseline	9.7		18.5		1.7	
	Week 8	5.2	−4.5	5.6	−12.9	1.2	−0.5
	Week 16	5.1	−4.6	4.6	−13.9	1.5	−0.2
4	Baseline	9.1		10.7		2.3	
	Week 8	8.2	−0.9	7.6	−3.1	1.8	−0.5
	Week 16						
8	Baseline	10.4		8.8		3	
	Week 8						
	Week 16	6.1	−4.3	3.8	−5	1.7	−1.3
11	Baseline	9.4		5.7		0.9	
	Week 8	10.2	0.8	5.2	−0.5	0.9	0
	Week 16						
12	Baseline	9.2		13.5		1.6	
	Week 8	8.2	−1	12.3	−1.2	2.5	0.9
	Week 16						
14	Baseline	8.8		8		2	
	Week 8	7.7	−1.1	2.7	−5.3	2.2	0.2
	Week 16						
Average (SD) (week 8)	n	6		6		6	
	Baseline	9.0 (0.7)		10.8 (4.6)		1.7 (0.5)	
	Week 8	7.6 (1.6)	−1.3 (1.6)	6.5 (3.0)	−4.3 (4.2)	1.7 (0.6)	0.0 (0.5)
Average (SD) (week 16)	n	3		3		3	
	Baseline	9.2 (1.5)		11.8 (5.8)		2.1 (0.8)	
	Week 16	5.8 (0.6)	−3.4 (1.6)	4.0 (0.6)	−7.8 (5.3)	1.7 (0.2)	−0.4 (0.8)

[a]HbA1c: hemoglobin A1c.

[b]FPG: fasting plasma glucose.

[c]TG: triglycerides.

[d]SD: standard deviation.

Seven-point glucose measurements at baseline and at visits 8 and 16 are shown in Figure 2. Patients' SF-12 and daily life questionnaires were similar between baseline and subsequent visits (Multimedia Appendix 1). Data regarding the cumulative time the mobile app was used was not available due to a technical malfunction, yet, most patients reported it was easy for them to adhere to the program (Multimedia Appendix 1).

Discussion

Principal Findings

Our study demonstrates an improvement in weight, BP, and glycemic parameters of patients provided with a stress reducing mobile app, in addition to standard of care.

The use of biofeedback-assisted relaxation in patients with T2D has been previously described. Mcginnis et al randomized 39 patients to a regimen of either 10 sessions of biofeedback and relaxation or 3 sessions of education. Biofeedback and relaxation were associated with significant decreases in average blood glucose, HbA1c, and muscle tension compared with the control group [9].

Jyotsna et al randomized 120 patients to either standard diabetes care or to a yogic breathing program for 6 months. At 6 months, quality of life and postprandial plasma glucose significantly improved in the group practicing yoga compared with baseline, but there was no significant improvement in the FPG and in

HbA1c [12]. Biofeedback has also been shown to reduce food craving [13] as well as BP [14,15].

There may be multiple mechanisms whereby biofeedback-assisted relaxation or yogic breathing may improve glycemic control. Activation of the sympathetic system during stress increases cortisol levels and insulin resistance. Biofeedback has been shown to decrease multiple aspects of the chronic physiological stress response including muscle tension, peripheral vasoconstriction, heart rate, cortisol, and cathecolamines. Biofeedback may also carry psychological effects such as an improved sense of control and better problem solving [8]. Diabetes control is largely dependent on the compliance of the patient with the drug and dietary regimen prescribed and patient empowerment is an inseparable aspect of diabetes care [16]. Improved problem solving and sense of control as well as reduction of depression, anxiety, and improved sleep patterns secondary to biofeedback may also contribute to improved glycemic control [14].

Strength and Limitations

The benefit of our approach of incorporating a mobile app stress reducing technique is its scalability. The app can be delivered to thousands of patients without incurring the cost of face-to-face classes of yogic breathing or biofeedback. In the current era where the use of mobile devices is increasing rapidly, incorporating mobile apps in the motivational programs for patients with diabetes may be effectively cost reducing.

Several limitations to our report must be recognized. The study was designed as a pilot study, and the number of patients completing the trial was small. Additionally, the trial included both the use of the mobile app as well as dietary consultation and weekly phone contact. Although patients had long-standing diabetes and were already well educated in the concepts of healthy diabetes nutrition, there are still well-recognized benefits of a “refresher” course in medical nutrition therapy, particularly when paralleled by weekly motivational calls. Finally, due to technical reasons, data regarding the extent of use of the app was not available, therefore, it is not possible to assess whether the extent of clinical benefit correlated with the extent of app usage.

Nevertheless, the results of the study as well as the positive feedback received from our patients encourage expansion of the study to a larger-scale randomized trial, wherein patients are randomized to standard of care alone or with the addition of the mobile app.

Acknowledgments

We are grateful to the diabetes unit team and Dr Olga Foiering for their support of the trial.

Conflicts of Interest

MM received funding from Eco Fusion for conduct of the trial. SAK and OF are the medical and general directors of Eco Fusion. AC has no relevant disclosures.

Multimedia Appendix 1

Patient questionnaires.

[PDF File (Adobe PDF File), 64KB - mhealth_v5i5e75_app1.pdf]

References

1. Diabetes Atlas. Idf Diabetes Atlas - 7th edition URL: http://www.diabetesatlas.org/ [accessed 2017-01-24] [WebCite Cache ID 6nku2ZRpz]
2. Tyagi A, Cohen M, Reece J, Telles S, Jones L. Heart rate variability, flow, mood and mental stress during yoga practices in yoga practitioners, non-yoga practitioners and people with metabolic syndrome. Appl Psychophysiol Biofeedback 2016 Dec;41(4):381-393. [doi: 10.1007/s10484-016-9340-2] [Medline: 27457341]
3. Renzaho AM, Houng B, Oldroyd J, Nicholson JM, D'Esposito F, Oldenburg B. Stressful life events and the onset of chronic diseases among Australian adults: findings from a longitudinal survey. Eur J Public Health 2014 Feb;24(1):57-62. [doi: 10.1093/eurpub/ckt007] [Medline: 23397581]
4. Pouwer F, Kupper N, Adriaanse MC. Does emotional stress cause type 2 diabetes mellitus? a review from the European Depression in Diabetes (EDID) Research Consortium. Discov Med 2010 Feb;9(45):112-118 [FREE Full text] [Medline: 20193636]
5. Rook KS, August KJ, Choi S, Franks MM, Stephens MA. Emotional reactivity to daily stress, spousal emotional support, and fasting blood glucose among patients with type 2 diabetes. J Health Psychol 2015 May 07 Epub ahead of print. [doi: 10.1177/1359105315581064] [Medline: 25953086]
6. Hegde S, Adhikari P, Kotian S, Pinto V, D'Souza S, D'Souza V. Effect of 3-month yoga on oxidative stress in type 2 diabetes with or without complications: a controlled clinical trial. Diabetes Care Oct 2011;34(10):2208-2210.
7. Wagner JA, Bermudez-Millan A, Damio G, Segura-Perez S, Chhabra J, Vergara C, et al. A randomized, controlled trial of a stress management intervention for Latinos with type 2 diabetes delivered by community health workers: Outcomes for

psychological wellbeing, glycemic control, and cortisol. Diabetes Res Clin Pract 2016 Oct;120:162-170. [doi: 10.1016/j.diabres.2016.07.022] [Medline: 27568646]
8. Surwit RS, van Tilburg MA, Zucker N, McCaskill CC, Parekh P, Feinglos MN, et al. Stress management improves long-term glycemic control in type 2 diabetes. Diabetes Care 2002 Jan;25(1):30-34. [Medline: 11772897]
9. McGinnis RA, McGrady A, Cox SA, Grower-Dowling KA. Biofeedback-assisted relaxation in type 2 diabetes. Diabetes Care 2005 Sep;28(9):2145-2149. [Medline: 16123481]
10. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [Medline: 8628042]
11. Gregoski MJ, Mueller M, Vertegel A, Shaporev A, Jackson BB, Frenzel RM, et al. Development and validation of a smartphone heart rate acquisition application for health promotion and wellness telehealth applications. Int J Telemed Appl 2012;2012:696324 [FREE Full text] [doi: 10.1155/2012/696324] [Medline: 22272197]
12. Jyotsna VP, Ambekar S, Singla R, Joshi A, Dhawan A, Kumar N, et al. Cardiac autonomic function in patients with diabetes improves with practice of comprehensive yogic breathing program. Indian J Endocrinol Metab 2013 May;17(3):480-485 [FREE Full text] [doi: 10.4103/2230-8210.111645] [Medline: 23869306]
13. Meule A, Freund R, Skirde AK, Vögele C, Kübler A. Heart rate variability biofeedback reduces food cravings in high food cravers. Appl Psychophysiol Biofeedback 2012 Dec;37(4):241-251. [doi: 10.1007/s10484-012-9197-y] [Medline: 22688890]
14. Mourya M, Mahajan AS, Singh NP, Jain AK. Effect of slow- and fast-breathing exercises on autonomic functions in patients with essential hypertension. J Altern Complement Med 2009 Jul;15(7):711-717. [doi: 10.1089/acm.2008.0609] [Medline: 19534616]
15. McGrady A. The effects of biofeedback in diabetes and essential hypertension. Cleve Clin J Med 2010 Jul;77 Suppl 3:S68-S71 [FREE Full text] [doi: 10.3949/ccjm.77.s3.12] [Medline: 20622080]
16. American Diabetes Association. 4. Lifestyle Management. Diabetes Care 2017 Jan;40(Suppl 1):S33-S43. [doi: 10.2337/dc17-S007] [Medline: 27979891]

Abbreviations

BP: blood pressure
FPG: fasting plasma glucose
HbA1c: hemoglobin A1c
SD: standard deviation
T2D: type 2 diabetes
TG: triglycerides
SMBG: self-monitoring of blood glucose

Edited by G Eysenbach; submitted 25.01.17; peer-reviewed by E Roitman, J Clarke; comments to author 02.03.17; revised version received 23.03.17; accepted 09.04.17; published 29.05.17.

Please cite as:
Munster-Segev M, Fuerst O, Kaplan SA, Cahn A
Incorporation of a Stress Reducing Mobile App in the Care of Patients With Type 2 Diabetes: A Prospective Study
JMIR Mhealth Uhealth 2017;5(5):e75
URL: http://mhealth.jmir.org/2017/5/e75/
doi:10.2196/mhealth.7408
PMID:28554881

Original Paper

Mobile App Design, Development, and Publication for Adverse Drug Reaction Assessments of Causality, Severity, and Preventability

Muslimah Ithnin[1*], BSc, MSc; Mohd Dzulkhairi Mohd Rani[1*], MBBS, MCommMed; Zuraidah Abd Latif[2*], MMed (Paed), MD; Paveethra Kani[2*], BPharm; Asmalita Syaiful[1*], BSc, MSc; Khairun Nain Nor Aripin[1*], MBChB, MSc, PhD; Tengku Amatullah Madeehah Tengku Mohd[1*], MBChB, MSc, MPH

[1]Universiti Sains Islam Malaysia, Kuala Lumpur, Malaysia

[2]Hospital Ampang, Ampang, Malaysia

[*]all authors contributed equally

Corresponding Author:
Tengku Amatullah Madeehah Tengku Mohd, MBChB, MSc, MPH
Universiti Sains Islam Malaysia
Tingkat 13, Menara B, Persiaran MPAJ
Jalan Pandan Utama, Pandan Indah
Kuala Lumpur, 55100
Malaysia
Phone: 60 3 4289 2400
Fax: 60 3 4289 2477
Email: madeehah@usim.edu.my

Abstract

Background: Adverse drug reactions (ADRs) cause significant morbidity and mortality. Improved assessment of ADRs to identify the causal relationship, the severity, and the preventability will aid ADRs prevention or reduce patient burden.

Objective: The aim of this study was to develop mobile apps in assisting clinical decision in ADR assessments of causality, severity, and preventability using validated tools. The usability of the apps was assessed.

Methods: We designed mobile apps using validated assessment tools for ADRs. They are the Liverpool ADRs Causality Assessment Tool, Hartwig's Severity Assessment Scale, and the Modified Schumock and Thronton Preventability Scale. The apps were named "Adverse Drug ReactionCausality," "Adverse Drug ReactionSeverity," and "Adverse Drug RxnPreventability." A survey was conducted using the System Usability Scale (SUS) to assess the usability of the developed apps among health care professionals.

Results: These apps are available for download through Google Play Store for free since January 2015. From the survey, the mean SUS score was 70.9 based on 26 responses from the pediatric ward of Hospital Ampang, Malaysia.

Conclusions: The developed apps received an overall acceptable usability among health care professionals. The usage of these apps will improve detection, assessment, and avoidance of future ADRs. They will also contribute to future research on ADRs, thus increasing drug safety.

(JMIR Mhealth Uhealth 2017;5(5):e78) doi:10.2196/mhealth.6261

KEYWORDS

mobile applications; computer-assisted decision making; drug monitoring; pharmacovigilance; adverse drug reactions

Introduction

Adverse drug reactions (ADRs) cause significant mortality and morbidity in patients [1-5]. The World Health Organization (WHO) defines an ADR as a response to a drug that is noxious and unintended and occurs at doses normally used in man for prophylaxis, diagnosis, or therapy of disease, or for modification of physiological function [6]. Previous studies have shown that ADRs were the cause of 3% of all hospital admissions in the pediatric population and that 10% of children suffer an ADR while in hospital [7]. Similar numbers are seen in adult patients [8]. ADRs have been estimated to cause 3% of all deaths in the

general population and up to 5% of deaths in hospitalized patients [4].

Previous studies have shown that the combined use of mobile technology and mobile apps software in health care offer various benefits to many parties including health care professionals, patients, management, and even stakeholders [9-13]. Mobile technology facilitates efficient delivery of services to patients. It also improves quality and effectiveness of services to the benefit of patients [14,15]. However, the use of apps in health care services does not seem as extensive as compared with other services such as social or public services [10,16]. Research in this area is needed to keep up with the increasing number of ADRs and integrate new knowledge with the rapidly advancing technology [9].

According to the WHO, ADR causality assessment is "a method by which it estimates the relationship concerning the agent (which is the drug) and the adverse reactions" [17]. It assesses the causal connection between the drugs and their adverse effects. Assessment of ADRs causality will give an advantage in the ability to classify the relationship, improve scientific evaluation for each individual ADR, and thus enable for an early warning system for clinicians, pharmacists, and health regulators [18-20].

The term severity in ADRs is used to describe the intensity of the adverse drug reaction [21,22]. Similar to causality assessment, severity assessment of ADRs are also crucial in epidemiological studies. The ability to classify the severity of ADRs will provide a mechanism for the health care workers and authorities to identify the problem areas and improve the intervention for patient care that would reduce the burden of ADRs [7,23,24].

Assessment of preventability is important for ADRs as it gives important information to improve prescription practice and enhance patient monitoring [16,24]. Although the assessment of each ADR's causality, severity, and preventability is crucial to provide important drug safety information, relatively few of these assessments are being performed [7].

This study was designed to develop apps for the aforementioned ADR assessments using validated tools. The apps can be downloaded on a mobile phone or mobile devices, which are then adapted to improve knowledge on ADRs and ultimately drug safety in health care.

Methods

Assessment Tool for Adverse Drug Reaction (ADR) Causality, Severity, and Preventability

There is currently no operational tool that has been proven as a gold standard for each ADR assessment of causality, severity, and preventability; therefore, the most widely used or accepted operational tools were selected for the development of our apps. Each of these tools has been validated by previous studies [16,25-29].

The ADR causality app was developed using the Liverpool ADR causality assessment tool [28]. This is a questionnaire-based classification for suspected ADRs using an algorithm built by a multidisciplinary team from the University of Liverpool in 2012. The algorithm classifies the suspected ADRs as definite, probable, possible, or unlikely. Results from a systematic review on assessment of ADR causality showed that the Naranjo algorithm was the most frequently used tool [7]. However, the Liverpool ADR causality assessment tool showed full range of causality category and good interrater reliability (IRR) compared with Naranjo algorithm. Thus, this tool was used in developing the ADR causality app [28].

The ADR severity app was developed based on the Hartwig's Severity Assessment Scale [24], which is the most commonly used severity tool in ADR studies. It classifies the ADR into mild, moderate, or severe based on level of clinical outcomes [7].

The ADR preventability app was developed using the Modified Schumock and Thornton Preventability Scale [29], which is the most frequently used scale in ADR studies in children [7]. It is a questionnaire on the criteria for determining preventability of ADRs based on clinical circumstances surrounding the ADR. The category of preventability is either definitely preventable, probably preventable, or not preventable.

Development and Publishing the App Into Google Play Store

The apps were developed using the rapid application development (RAD) model [30]. Using this model, the development processes are divided into three main phases which are preproduction, production, and postproduction.

For the development of the ADR assessment apps, Windows 8.1 by Microsoft was used as the operational system. MIT App Inventor Tool version 2.3.0 [31] was used during the production process and aiStater emulator [32] was then used to provide communication between App Inventor running in the browser and other parts of App Inventor. App Inventor is a free, cloud-based service accessed with a Google account.

After all the production phases were completed, the app was then saved in an APK file and then uploaded into the Google Play Developer Console. Once the ADR assessment app was published in Google Play Store, it could then be downloaded and installed for free by Android OS users.

Table 1. Adverse drug reactions (ADRs) assessment tool.

Assessment tool	Reference	App name
Causality assessment	Liverpool Adverse Drug Reaction Causality Assessment Tool [28]	Adverse Drug ReactionCausality
Severity assessment	Hartwig's Severity Assessment Scale [24]	Adverse Drug ReactionSeverity
Preventability assessment	Modified Schumock and Thronton Preventability Scale [29]	Adverse Drug RxnPreventability

The apps were designed without storage capacity to avoid issues regarding patient confidentiality or personal data. Therefore, information input into the system is not available to anyone. The apps are also accessible for offline use. The details for each ADR app and the references used are shown in Table 1.

Testing and Measure of App Usability

The System Usability Scale (SUS), a reliable and low-cost usability scale, was used to assess the usability of the ADR app. SUS is a 10-item scale presented with a 5-point Likert scale, which results in an overall score from 0 to 100 that indicates the perceived usability of the interface [33].

A survey was conducted among 26 health care professionals in the pediatric ward of Hospital Ampang, Malaysia, where they were asked to answer the SUS questionnaire. The survey was conducted 10 months after the introduction of the apps among staff at the pediatric ward of Hospital Ampang.

Results of the SUS questionnaire were recorded and normalized using SPSS version 20 (IBM Corp). The mean SUS score and the standard deviation (SD) were then recorded. Products with scores <70 were considered candidates for increased scrutiny, and continued improvement was judged to be marginal at best [34].

Results

App Development and Publication

The developed apps were published in Google Play Store on January 22, 2015. All 3 apps were considered to have fulfilled the objective of the development. The apps were freely downloadable from Google Play Store from February 2015. Exemplar screenshots for each app are shown in Figure 1 for causality, Figure 2 for severity, and Figure 3 for preventability.

Up until January 20, 2017, a total of 609 users have downloaded the apps. The total installer, installer by country, and ranking statistics for each ADRs assessment app are shown in Table 2. The highest numbers of downloads were for the causality app followed by the severity and preventability apps. The installers were mainly from India and Malaysia for all the apps.

App Usability Among Health Care Professionals

Of the 26 health care professionals involved in the survey, 19 (73.1%) of the respondents were physicians, 6 respondents (23.1%) were nurses, and 1 respondent (3.8%) was a pharmacist. The mean SUS score was 70.9 (SD 12.86). The results showed that the SUS score was >70; thus, the app tested is within the acceptable range of usability [34]. Table 3 depicts the responses to the usability-related questions.

Figure 1. Exemplar of adverse drug reaction (ADR) causality assessment tool screenshot.

Figure 2. Exemplar of adverse drug reaction (ADR) severity assessment tool screenshot.

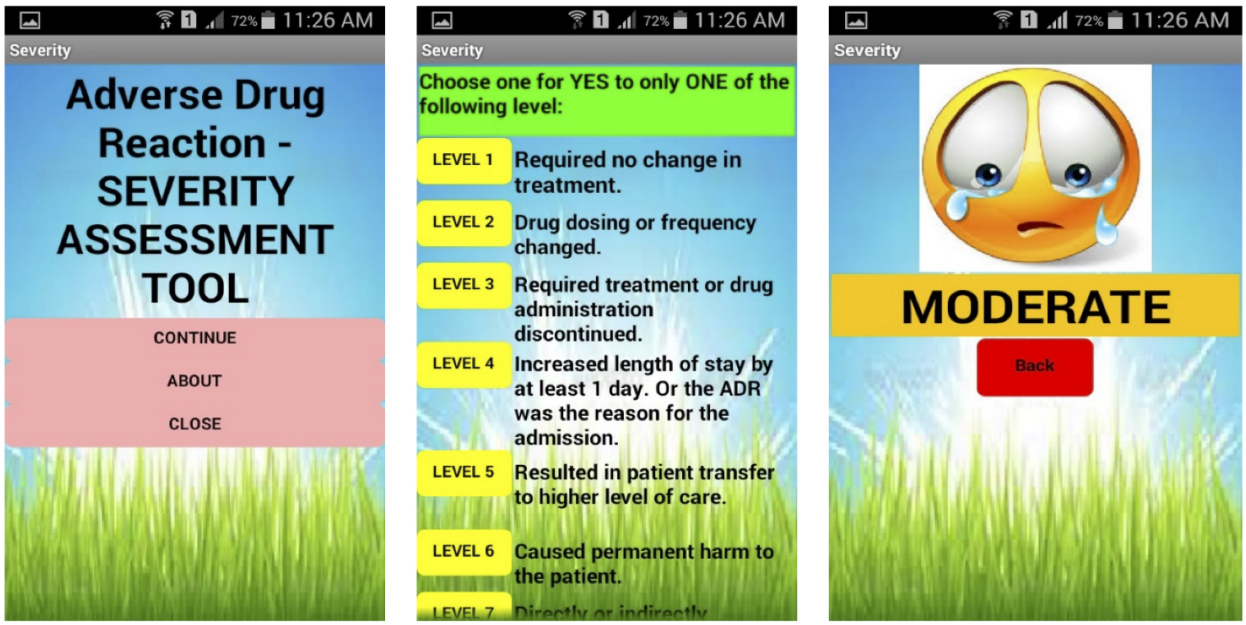


Figure 3. Exemplar of adverse drug reaction (ADR) preventability assessment tool screenshot.

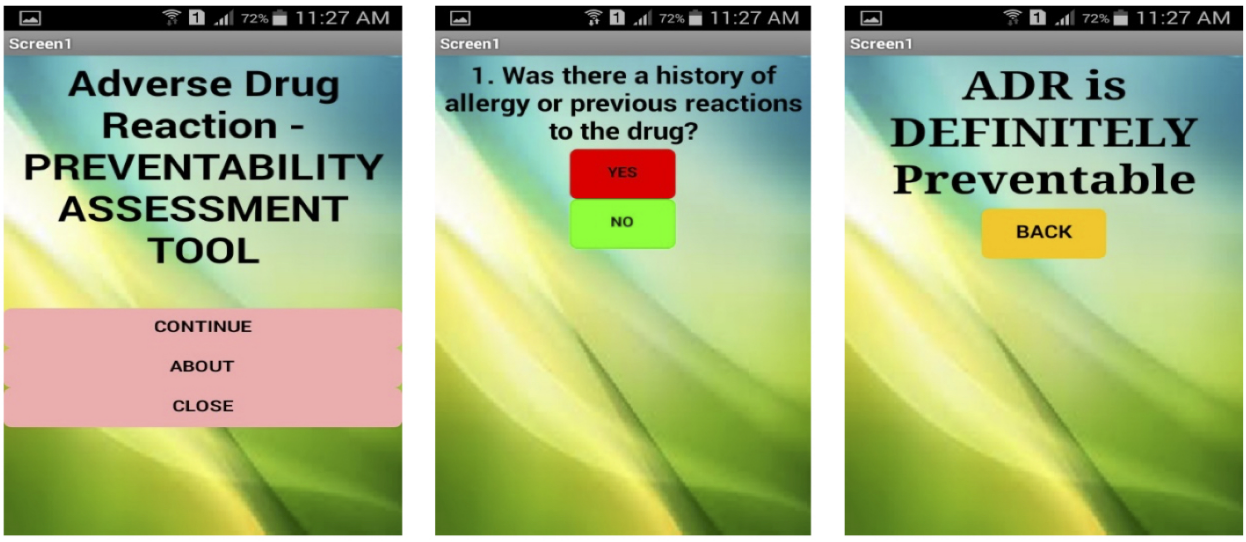


Table 2. Statistics of adverse drug reaction (ADR) assessment apps from Google Play Store.

Characteristics	Adverse Drug ReactionCausality	Adverse Drug ReactionSeverity	Adverse Drug RxnPrevenrability
Total installers	274	200	135
Installer by country (%)	India 77 (28.1)	Malaysia 70 (35)	India 48 (35.5)
	Malaysia 69 (25.0)	India 61 (30.5)	Malaysia 31 (23.0)
	Saudi Arabia 43 (15.6)	Qatar 17 (8.5)	Qatar 16 (11.9)
	South Africa 17 (6.3)	Saudi Arabia 17 (8.5)	Others 40 (29.6)
	United States 17 (6.3)	South Africa 17 (8.5)	
	Others 51 (18.7)	Others 18 (9)	
Average ratings	4.75	5.00	4.50

Table 3. Usability questions and summary of responses (N=26).

Answer option	Strongly disagree n (%)	Disagree n (%)	Neutral n (%)	Agree n (%)	Strongly agree n (%)
I think that I would like to use this system frequently	0 (0)	0 (0)	3 (11)	21 (81)	2 (8)
I found the system unnecessarily complex	5 (19)	10 (38)	8 (31)	3 (12)	0 (0)
I thought the system was easy to use	0 (0)	1 (4)	1 (4)	18 (69)	6 (23)
I think that I would need the support of a technical person to be able to use this system	6 (23)	11 (42)	3 (12)	6 (23)	0 (0)
I found the various functions in this system were well integrated	1 (4)	2 (8)	7 (27)	14 (53)	2 (8)
I thought there was too much inconsistency in this system	3 (12)	12 (45)	9 (35)	2 (8)	0 (0)
I would imagine that most people would learn to use this system very quickly	0 (0)	0 (0)	4 (15)	14 (54)	8 (31)
I found the system very cumbersome to use	5 (19)	12 (46)	5 (19)	4 (16)	0 (0)
I felt very confident using the system	0 (0)	0 (0)	7 (27)	16 (61)	3 (12)
I needed to learn a lot of things before I could get going with this system	5 (19)	15 (57)	3 (12)	3 (12)	0 (0)

Discussion

Principal Findings

The causality app had the highest number of installers so far. This seems to be similar with previous 102 published ADR studies where causality was the most common assessment conducted in suspected ADR cases [7].

Based on the country of origin, the highest percentage of installers were from India for all of the ADR apps published. India is currently working to strengthen its pharmacovigilance program due to the rapidly growing number of ADR studies in the country [35-38]. We expect that our ADR apps would be able to assist not only Indian researchers and clinical researchers but also any center conducting research on drug safety.

The apps usability among health professionals in the hospital was assessed using SUS, which consisted of 10 alternate statements of positive and negative items rated using a 5-point Likert scale. Our survey results show that the apps developed have a mean SUS of 70.9 (SD 12.86), thus demonstrating acceptable usability.

The health care professionals that used the apps concluded that the apps were convenient and they would choose them over conventional paper-based assessments. Previous studies have found that the use of apps in health care is cost-effective, faster, easier, and more interactive due to factors of mobility, convenience, and involvement of active touching of the screen to perform the assessment [11,39]. The apps are also secure, as they do not store any information from the data inputted into the app [40].

The use of medical apps by health care professionals and researchers, and the numbers of these apps are increasing rapidly. Apps can give additional advantages at the point of care such as in diagnosis, monitoring, reporting, or follow-up of treatment [12]. The increased usage of mobile phone or mobile device apps warrant further studies evaluating their utility and effectiveness on a larger scale.

Limitations

We have identified a few limitations of the apps. On a basic level, there is no assessment tool universally accepted or described as the gold standard for ADRs either for causality, severity, or preventability. We chose the most widely used and validated algorithms and scales to develop the apps; however, we recognize that not all researchers will agree with the algorithms chosen in development of ADR assessment.

Second, the apps have only been evaluated by health care professionals from the pediatric department in a hospital setting. Further evaluation is necessary to gain more feedback from a wider range of users.

Finally, the aesthetics of the app contents in terms of color, text letters, and pictures have been optimized; however, there is room for improvement to make the apps more attractive.

We are continuously working to update and upgrade the apps. Future research is needed to test the usability of the apps in varying populations and to add several other commonly used algorithms or tools in ADRs assessment. Research to highlight the context and content of the apps should also be designed specifically for health care professionals, researchers, and regulators.

Conclusions

These ADR assessment apps will aid health care professionals in determining the causality, severity, and preventability of ADRs. This is aimed to contribute toward efforts to reduce the burden of ADRs on patients. The SUS score data showed that the apps have acceptance usability among health care professionals. They will also support future research to enhance overall safety relating to drugs given to patients.

Acknowledgments

We would like to thank all the doctors, nurses, and pharmacists from Hospital Ampang who were involved in the survey. A special thanks to Mr Che Mat bin Muda for helping us to program the apps. This study was funded by Malaysian Ministry of Higher Education (MOHE), grant no. USIM/RAGS/FPSK/36/50314. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

KN, MDMR, and TAMTM conceived and supervised the study. MI and AS designed and published the app. They, together with ZAL and PK, tested the apps. MI wrote the paper, while KN and TAMTM edited the manuscript and gave the final approval. All authors contributed exclusively to the work presented in this paper.

Conflicts of Interest

None declared.

References

1. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998;279(15):1200-1205. [Medline: 9555760]
2. Impicciatore P, Choonara I, Clarkson A, Provasi D, Pandolfini C, Bonati M. Incidence of adverse drug reactions in paediatric in/out - patients: a systematic review and meta - analysis of prospective studies. Brit J Clin Pharmaco 2001;52(1):77-83. [doi: 10.1046/j.0306-5251.2001.01407.x]
3. Le J, Nguyen T, Law AV, Hodding J. Adverse drug reactions among children over a 10-year period. Pediatrics 2006;118(2):555-562. [doi: 10.1542/peds.2005-2429] [Medline: 16882807]
4. Wester K, Jönsson AK, Spigset O, Druid H, Hägg S. Incidence of fatal adverse drug reactions: a population based study. Brit J Clin Pharmaco 2008;65(4):573-579. [doi: 10.1111/j.1365-2125.2007.03064.x]
5. Kongkaew C, Noyce PR, Ashcroft DM. Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. Ann Pharmacother 2008;42(7-8):1017-1025.
6. World Health Organization. International drug monitoring. The role of the hospital. Geneva: World Health Organization; 1969.
7. Smyth RM, Gargon E, Kirkham J, Cresswell L, Golder S, Smyth R, et al. Adverse drug reactions in children—a systematic review. PLoS One 2012;7(3):e24061. [doi: 10.1371/journal.pone.0024061]
8. Khan LM. Comparative epidemiology of hospital-acquired adverse drug reactions in adults and children and their impact on cost and hospital stay - a systematic review. Eur J Clin Pharmacol 2013;69(12):1985-1996. [doi: 10.1007/s00228-013-1563-z]
9. Hameed K. The application of mobile computing and technology to health care services. Telemat Inform 2003;20(2):99-106 [FREE Full text]
10. Dehlingher J, Dixon J. Mobileseworkshop. 2011. Mobile application software engineering: challenges and research directions URL: http://www.mobileseworkshop.org/papers/7_Dehlinger_Dixon.pdf [accessed 2017-05-26] [WebCite Cache ID 6qjxAzB2s]
11. Holzinger A, Kosec P, Schwantzer G, Debevc M, Hofmann-Wellenhof R, Frühauf J. Design and development of a mobile computer application to reengineer workflows in the hospital and the methodology to evaluate its effectiveness. J Biomed Inform 2011;44(6):968-977. [doi: 10.1016/j.jbi.2011.07.003]
12. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis 2012;12(1):67 [FREE Full text] [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]
13. Ahtinen A, Mattila E, Välkkynen P, Kaipainen K, Vanhala T, Ermes M. Mobile mental wellness training for stress management: feasibility and design implications based on a one-month field study. JMIR mHealth and uHealth 2013;1(2):e11. [doi: 10.2196/mhealth.2596] [Medline: 25100683]
14. Lu Y, Xiao Y, Sears A, Jacko JA. A review and a framework of handheld computer adoption in healthcare. Int J Med Inform 2005 Jun;74(5):409-422. [doi: 10.1016/j.ijmedinf.2005.03.001] [Medline: 15893264]
15. Holzinger A, Searle G, Kleinberger T, Seffah A, Javahery H. Investigating usability metrics for the design and development of applications for the elderly. In: Miesenberger K, Klaus J, Zagler W, Karshmer A, editors. Computers Helping People with Special Needs. Berlin, Heidelberg: Springer; 2008:98-105.
16. Holzinger A, Errath M. Mobile computer web-application design in medicine: some research based guidelines. Universal Access Inf 2007;6(1):31-41 [FREE Full text]
17. World Health Organization. WHO. Uppsala: The Uppsala Monitoring Centre; 2005. The use of the WHO-UMC system for standardized case causality assessment URL: https://www.who-umc.org/media/2768/standardised-case-causality-assessment.pdf [accessed 2014-01-25] [WebCite Cache ID 6qWqnh5Ve]
18. Meyboom RH, Hekster YA, Egberts AC, Gribnau FW, Edwards IR. Causal or casual? The role of causality assessment in pharmacovigilance. Drug Saf 1997;17(6):374-389. [Medline: 9429837]

19. Arimone Y, Bégaud B, Miremont-Salamé G, Fourrier-Réglat A, Moore N, Molimard M, et al. Agreement of expert judgment in causality assessment of adverse drug reactions. Eur J Clin Pharmacol 2005;61(3):169-173. [doi: 10.1007/s00228-004-0869-2]
20. Macedo AF, Marques FB, Ribeiro CF, Teixeira F. Causality assessment of adverse drug reactions: comparison of the results obtained from published decisional algorithms and from the evaluations of an expert panel. Pharmacoepidem Dr S 2005;14(12):885-890. [doi: 10.1002/pds.1138]
21. ICH Expert Working Group. Clinical safety data management: definitions and standards for expedited reporting ICH-2 EWG E2. 1994 Presented at: International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use; October 27, 1994; Geneva, Switzerland URL: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf
22. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. The Lancet 2000;356(9237):1255-1259. [doi: 10.1016/S0140-6736(00)02799-9] [Medline: 11072960]
23. Doherty MJ. Algorithms for assessing the probability of an adverse drug reaction. Respiratory Medicine CME 2009;2(2):63-67 [FREE Full text] [doi: 10.1016/j.rmedc.2009.01.004]
24. Hartwig SC, Siegel J, Schneider PJ. Preventability and severity assessment in reporting adverse drug reactions. Am J Hosp Pharm 1992;49(9):2229-2232. [Medline: 1524068]
25. Agbabiaka TB, Savović J, Ernst E. Methods for causality assessment of adverse drug reactions. Drug Safety 2008;31(1):21-37. [doi: 10.2165/00002018-200831010-00003]
26. Ferner RE, Aronson JK. Preventability of drug-related harms—part I. Drug Safety 2010;33(11):985-994. [doi: 10.2165/11538270-000000000-00000]
27. Benkirane R, Soulaymani-Bencheikh R, Khattabi A, Benabdallah G, Alj L, Sefiani H, et al. Assessment of a new instrument for detecting preventable adverse drug reactions. Drug Safety 2015;38(4):383-393. [doi: 10.1007/s40264-014-0257-5]
28. Gallagher RM, Kirkham JJ, Mason JR, Bird KA, Williamson PR, Nunn AJ, et al. Development and inter-rater reliability of the Liverpool adverse drug reaction causality assessment tool. PLoS One 2011;6(12):e28096. [doi: 10.1371/journal.pone.0028096]
29. Mrugank BP, Hareesha RP. Prospective observational, non-randomized, parallel sequence study for assessment of adverse drug reactions due to chemotherapeutic treatment in different types of cancer patients. Int J Pharm Sci Res 2013;4(1):386-391.
30. Martin J. Rapid application development. Indianapolis, IN: Macmillan Publishing Co; 1991.
31. Walter D, Sherman M. Learning MIT App Inventor: A Hands-on Guide to Building Your Own Android Apps. Crowfordsville, IN: Pearson Education; 2014.
32. Massachusetts Institute of Technology. MIT App Inventor. 2012. Installing and running the emulator in AI2 URL: http://appinventor.mit.edu/explore/ai2/setup-emulator.html [accessed 2014-01-15] [WebCite Cache ID 6qWtReeme]
33. Bangor A, Kortum P, Miller JA. The system usability scale (SUS): an empirical evaluation. Int J Hum Comput Interact 2008;24(6):574-594.
34. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Studies 2009;4(3):114-123.
35. Priyadharsini R, Surendiran A, Adithan C, Sreenivasan S, Sahoo FK. Study of adverse drug reactions in pediatric patients. J Pharmacol Pharmacother 2011;2(4):277-280.
36. Digra KK, Pandita A, Saini GS, Bharti R. Pattern of adverse drug reactions in children attending the department of pediatrics in a tertiary care center: a prospective observational study. Clin Med Insights Pediatr 2015;9:73-78. [doi: 10.4137/CMPed.S29493]
37. Ponnusankar S, Tejaswini M, Chaitanya M. Assessment of Adverse Drug Reactions Based on Spontaneous Signals at Secondary Care Public Hospital. Indian J Pharm Sci 2015;77(4):490-493 [FREE Full text] [Medline: 26664067]
38. Kalaiselvan V, Kaur I, Singh S, Singh GN. Pharmacovigilance Programme of India: system put in place to report adverse drug reactions. Indian J Pharm Educ 2015;50(1):212-214 [FREE Full text] [doi: 10.5530/ijper.50.1.26]
39. Weaver A, Young AM, Rowntree J, Townsend N, Pearson S, Smith J, et al. Application of mobile phone technology for managing chemotherapy-associated side-effects. Ann Oncol 2007;18(11):1887-1892. [doi: 10.1093/annonc/mdm354]
40. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015;6:27882. [doi: 10.3402/ejpt.v6.27882]

Edited by G Eysenbach; submitted 27.06.16; peer-reviewed by L Sheets, S Sahay, F Sanz, S Abideen, R Altman; comments to author 18.11.16; revised version received 27.01.17; accepted 15.03.17; published 30.05.17.

Please cite as:
Ithnin M, Mohd Rani MD, Abd Latif Z, Kani P, Syaiful A, Nor Aripin KN, Tengku Mohd TAM
Mobile App Design, Development, and Publication for Adverse Drug Reaction Assessments of Causality, Severity, and Preventability
JMIR Mhealth Uhealth 2017;5(5):e78
URL: http://mhealth.jmir.org/2017/5/e78/
doi:10.2196/mhealth.6261
PMID:28559222

Original Paper

Mobile Technology Interventions for Asthma Self-Management: Systematic Review and Meta-Analysis

Lisa Miller[1], M Psych; Benjamin Schüz[1], PhD; Julia Walters[2], PhD; E Haydn Walters[2], PhD

[1]School of Medicine, Psychology, University of Tasmania, Hobart, Australia

[2]University of Tasmania, Hobart, Australia

Corresponding Author:
Lisa Miller, M Psych
School of Medicine, Psychology
University of Tasmania
Private Bag 30
Hobart, TAS 7001
Australia
Phone: 61 3 6226 7471
Fax: 61 3 6226 2883
Email: lisa.miller@utas.edu.au

Abstract

Background: Mobile technology interventions (MTI) are becoming increasingly popular in the management of chronic health behaviors. Most MTI allow individuals to monitor medication use, record symptoms, or store and activate disease-management action plans. Therefore, MTI may have the potential to improve low adherence to medication and action plans for individuals with asthma, which is associated with poor clinical outcomes.

Objective: A systematic review and meta-analysis were conducted to evaluate the efficacy of MTI on clinical outcomes as well as adherence in individuals with asthma. As the use of evidence-based behavior change techniques (BCT) has been shown to improve intervention effects, we also conducted exploratory analyses to determine the role of BCT and engagement with MTI as moderators of MTI efficacy.

Methods: We searched electronic databases for randomized controlled trials up until June 2016. Random effect models were used to assess the effect of MTI on clinical outcomes as well as adherence to preventer medication or symptom monitoring. Mixed effects models assessed whether the features of the MTI (ie, use of BCT) and how often a person engaged with MTI moderated the effects of MTI.

Results: The literature search located 11 studies meeting the inclusion criteria, with 9 providing satisfactory data for meta-analysis. Compared with standard treatment, MTI had moderate to large effect sizes (Hedges *g*) on medication adherence and clinical outcomes. MTI had no additional effects on adherence or clinical outcomes when compared with paper-based monitoring. No moderator effects were found, and the number of studies was small. A narrative review of the two studies, which are not included in the meta-analysis, found similar results.

Conclusions: This review indicated the efficacy of MTI for self-management in individuals with asthma and also indicated that MTI appears to be as efficacious as paper-based monitoring. This review also suggested a need for robust studies to examine the effects of BCT use and engagement on MTI efficacy to inform the evidence base for MTI in individuals with asthma.

(JMIR Mhealth Uhealth 2017;5(5):e57) doi:10.2196/mhealth.7168

KEYWORDS

asthma; mhealth; medication adherence; patient monitoring; behavior and behavior mechanisms; meta-analysis

Introduction

Background

Asthma is a chronic inflammatory disease of the airways that is characterized physiologically by excessive variation in airflow and manifests symptomatically as repeated episodes of coughing, wheezing, shortness of breath, and chest tightness [1]. Asthma affects 2.5 million Australians (11%) across the lifespan [2].

In terms of treatment recommendations, the National Asthma Council Australia and Global Strategy for Asthma Management and Prevention have recommended the use of personalized action plans and daily preventer medications to manage the illness [1,3]. Action plans are associated with improved clinical outcomes through incorporation of appropriate education, self-monitoring of symptoms and medication use, and review of symptoms by the individual between physician visits [4]. In more severe patients, daily preventer treatment is associated with a reduced risk of death, as well as asthma exacerbations, thus requiring hospitalization or oral steroid with their related side effects [5,6].

Despite current guidelines recommending action plans and daily preventer medication, patients' actual adherence to either treatment remains low. In 2015, only 28.10% (702,500/2,500,000) of Australians with asthma had an action plan [2]. Similarly, medication adherence is also suboptimal, with only 59.63% (1601/2685) of those with asthma reported using preventer medication at least once within a 12-month period, and only 33.89% (910/2685) reporting daily use [6]. Objective data on the dispensing of prescriptions for asthma preventer medication indicate that less than 18.00% (253,123/1,406,240) of those with asthma use preventer medication daily [7]. A longitudinal study following middle-aged adults with asthma over 12 months found that 73.8% (259/351) used inadequate preventer medication [8]. This is concerning, as research suggests that appropriate use of preventer medication could protect against progressive decline in lung function, which is associated with increasing asthma severity [8].

Lower adherence is associated with individuals who perceive their asthma to be an acute condition with few adverse personal consequences and thus has a lower perceived necessity for the preventer treatment [9]. The Common Sense Model of Illness [10] provides a theoretical explanation for asthma nonadherence. Information provided to the individual should inform perceptions of the causes, consequences, controllability, identity, and timeline of asthma, which drive coping responses around adherence, management, and so on. This in turn should as a consequence have positive effects on clinical outcomes, that is, exacerbation rates, hospitalizations, and so on [10].

Mobile Technology Interventions

Mobile technology interventions (MTI) can provide an external source of self-management support, allow for accurate real-time symptom and medication monitoring, and use built-in reminders to adhere to treatment, as well as can store and activate action plans. Studies suggest that over 80% of people with asthma are willing to use MTI, and quantitative studies suggest that it is an acceptable medium for assisting asthma self-management [11,12]. This might be facilitated by MTI providing a feeling of support that positively influences individuals' ability to cope with their asthma [13], for example, by identifying asthma-related stressors [11]. It is also suggested that MTI may improve the quality of care and asthma-management skills as well as allow for greater understanding of client attitudes, interpretation, and misconceptions of their asthma management [14]. Further, MTI have shown positive short-term behavioral outcomes for self-management medication in diabetes and for antiretroviral therapy [15]. However, physicians have highlighted major concerns to MTI use including increased time and resource demands, accuracy of information, liability, and patient confidentiality [14].

Research on behavioral interventions suggests that emphasizing specific behavior change techniques (BCT) in MTI allows clearly defined objectives for intervention development and therefore improved replication [16]. A recent review of mobile intervention applications for medication adherence found that action plans, prompt or cues, self-monitoring, and feedback on behavior were the most commonly used BCT, but at the same time that most recent mobile applications made limited or no use of BCT [17]. Using BCT in MTI has several advantages: As observable working components of MTI, BCT allow for replication and creation of a sustainable evidence base—basically, it becomes easier to compare MTI efficacy based on their contents and develop more effective MTI on this evidence base. Further, the theoretical underpinning of BCT allows understanding and identifying the mechanisms of change in adherence behavior and informs future improved MTI development [16]. BCT can therefore increase the likelihood of MTI efficacy to improve medication adherence [17], in particular, if they are being deployed within a well-functioning health care provider-patient relationship [11], and the BCT align with the therapeutic approach to self-management. Currently, it is unclear to what degree MTI for asthma management apply evidence-based BCT or provide evidence-based content.

Behaviour change techniques, including action plans, self-monitoring, and feedback on behavior, may be unsuccessful unless the individual engages with the MTI. Engagement with the MTI is in itself a quantifiable measure, ie, how often the individual sends or responds to the MTI [18], and indeed the efficacy of the MTI is likely to be dependent on the active engagement of the individual [19]. However, the reporting of such engagement with MTI has been poor, and engagement appears to be widely variable [18].

MTI in Asthma: The Present Review

There have been two reviews of MTI in asthma self-management. Belisario and colleagues [20] conducted a systematic review and meta-analysis of two MTI of asthma self-management and found insufficient evidence to draw a conclusion. Tran and colleagues [21] conducted a systematic review of 6 MTI of asthma self-management and found MTI was associated with greater medication adherence compared with standard treatment, but this did not translate to improvements in quality of life, symptom control, or lung function. These studies highlighted the need to identify studies that included evidence-based techniques of MTI, which successfully instigated changes and sustainability of self-management behavior.

Aims of This Review

This study therefore aimed to investigate the efficacy of MTI for adherence (medication and self-monitoring) and clinical health outcomes (lung function, quality of life, asthma control, and unscheduled visits) in individuals with asthma. In addition, we conducted explanatory analyses to answer the following

important theoretical and applied research questions: (1) Is MTI more efficacious than standard treatment or paper-based monitoring? (2) What behaviour change techniques are effective via mobile technology interventionsI? (3) Does engagement with the MTI enhance outcomes?

Methods

Literature Search, Inclusion Criteria, and Study Selection

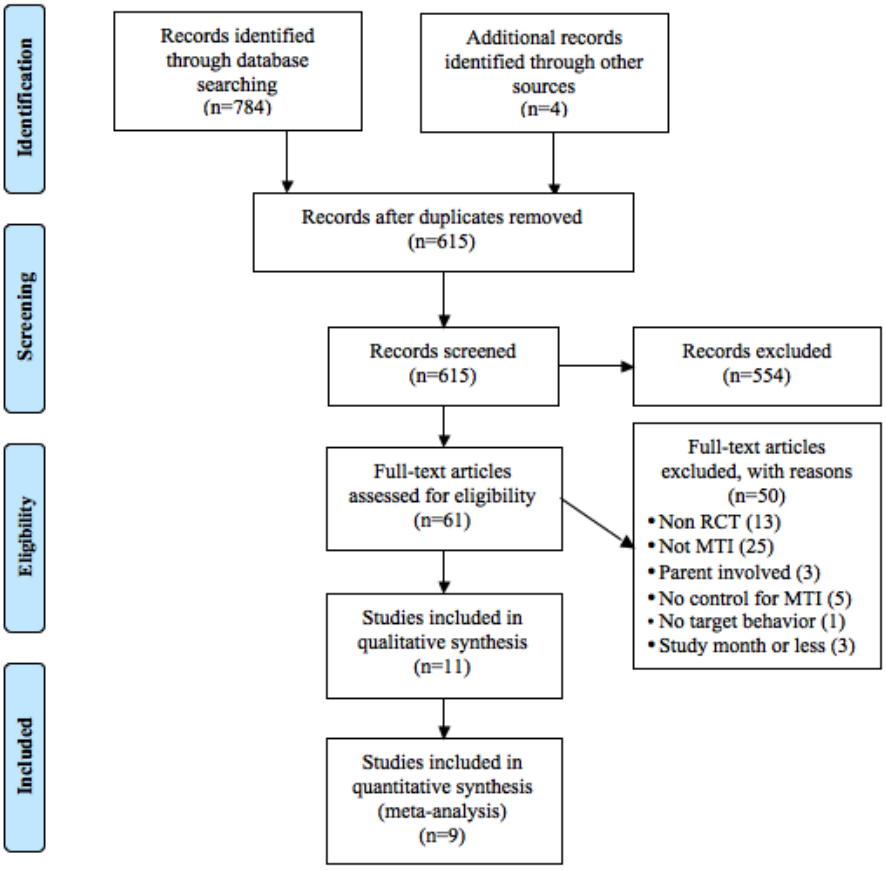
This review adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, [22]). A systematic literature review was conducted using Medical Literature Analysis and Retrieval System Online (MEDLINE), PsychNet, Scopus, and Web of Science to collect published studies as well as ProQuest Dissertations and Theses Global and ClinicalTrials.gov for relevant trials or unpublished studies. In addition, the reference lists of included studies and applicable systematic reviews were hand searched to identify additional studies. Titles, abstracts, and keywords were used to identify relevant studies (refer to Multimedia Appendix 1 for the full search strategy).

We included any randomized controlled trial using an MTI where the primary target of the MTI was the individual with asthma. Therefore, trials with children were excluded if the parent was responsible for asthma management. MTI were included if they used any mobile device that was currently available (ie, smart phone, tablet, or mobile phone) and any platform of delivery that incorporated at least one target behavior (ie, text message, application, music file). Papers not written in English were excluded. Screening individuals with asthma can result in improved asthma control over the proceeding week [23]. Therefore, to ensure that any changes in outcomes were due to the MTI, only the papers that exceeded the study duration of 1 month were included. The search was completed on the June 30, 2016, with the identification of 61 potential studies. Full texts of all eligible studies were retrieved and assessed against the inclusion criteria by LM and reviewed through discussion with BS. There were 11 studies included in the systematic review, with 9 included in the meta - analysis, on the basis of sufficient data. A flowchart of the study selection process can be found in Figure 1. The studies excluded from the full-text review and their reasons for omission can be found in Multimedia Appendix 2.

Study Characteristics

For each study, we extracted publication characteristics (eg, year of publication), the study design (eg, type of intervention, MTI characteristics), the sample (eg, female percentage, mean age, asthma severity), and the outcome variables (eg, adherence, symptom monitoring, lung function, quality of life) at pre- and postintervention or at change between pre- and postintervention time points. Data from the end point were taken for those studies reporting data over multiple times. A complete list of the extracted characteristics from all the identified studies included in the meta-analysis and the narrative review can be found in Multimedia Appendix 3.

Figure 1. Flowchart of the study selection process for the systematic review and meta-analysis.



Classification of BCT

None of the included studies specifically mentioned any behaviour change technique. We used the behavior change technology taxonomy [16] to code the features of the MTI. This taxonomy identifies 93 distinct and evidenced-based techniques within 16 categories of behavior-change interventions. The features of any paper-based monitoring and standard treatment groups were also classified according to the taxonomy for each of the included studies.

Risk of Bias Assessment

The quality of the included manuscripts was assessed according to the Cochrane Risk of Bias Guidelines [24]. The Cochrane Risk of Bias dimensions included random sequence generation, allocation concealment, blinding of outcome assessment, incomplete outcome data, selective reporting, and other bias [24]. The option of "blinding participants and personnel" was excluded because participants could not be blinded as to whether they did or did not receive the MTI. LM assessed the risk of bias in the included studies, which was reviewed through discussion with BS. Each domain was judged as "low," "high," or "unclear."

Primary and Secondary Outcomes

The primary outcomes were clinical health markers (ie, lung function, quality of life, asthma control, and unscheduled visits to the general physician or the emergency department) as well as frequency of self-monitoring and medication behaviors (adherence). Secondary outcomes were the type and number of behaviour change technique used and the frequency of mobile technology intervention use (engagement) .

Meta-Analytic Strategy

The most frequently reported effects were postintervention unstandardized means and standard deviations, or change from, and therefore Hedges *g* was the effect measure used for the meta‐analysis. If available, change data were used in preference to postintervention data, as change scores provide a more powerful estimate of the intervention effect [24]. Bias-corrected Hedges *g* was reported to account for the positively biased estimate of an effect size when sample sizes are small, which was common in the included studies [25].

The meta-analytic procedure was conducted using the *R* "metafor" package [26] applying a random effects model on pooled effect sizes of primary and secondary outcomes. The assumption of the random effects model is that variability between effect sizes is due to real between-study differences and allows for generalizability of findings beyond the included studies [27]. It was hypothesized that studies in the meta-analysis would provide multiple effect sizes because different outcome variables were being assessed. The magnitude of the effect was interpreted as small (0.20), moderate (0.50), and large (0.80) [25].

Heterogeneity of effect sizes was measured using *Q* and I^2statistics. Caution should be taken when interpreting these statistics as they are often imprecisely estimated in the presence of a small number of studies, as in this meta-analysis [24]. Funnel plots were visually examined to assess for publication bias. No tests of funnel plot asymmetry were conducted due to the small power of the test in a small sample of studies [24].

Mixed-effects meta-analysis was used to perform the categorical and continuous moderation analyses. Due to overlap in BCT between MTI and paper-based monitoring, only studies comparing MTI with standard treatment were included in the BCT moderation analyses. BCT were aggregated to 4 categorical variables (self-monitoring with feedback, self-monitoring only, prompting only, other) and were entered as categorical moderators. All relevant behavior change technique taxonomies (BCTT), with the exception of "Instruction on how to perform the behavior" (BCTT 4.1) and "Information about health consequences" (BCTT 5.1), contributed to the number of BCTT in the MTI group. Behavior change technique taxonomies 4.1 and 5.1 were excluded as they were also commonly used in the standard treatment group.

A categorical moderation analysis was used to assess the within- and between-group variability of effect sizes using a dummy-coded categorical factor to indicate the moderators. Categorical moderation is indicated by a significant between-group statistic (*QM*).

The numbers of BCT used in MTI and engagement with MTI were entered as continuous moderators for the effect size estimates, and the standardized regression coefficient of the moderator was examined for statistical significance. In addition, the R^2statistic indicates how much of the variation between studies can be explained by the model containing the moderator, relative to the total variation.

Results

Study Characteristics

Table 1 shows the population characteristics of the included studies. The mean sample size of 11 included studies was 103, ranging from 16 to 288. The mean study duration was 17 weeks, ranging from 12 to 24 weeks. The age of participants ranged from 10 years to about 65 years, with a mean age of 34 years, and the percentage of female participants was 61.1% (583/954). Asthma severity varied from mild to severe persistent asthma, but the method used to assess this was reported only in 5 studies [28-32].

The effect sizes for MTI were computed from comparisons of MTI with paper-based and standard treatment. Three studies compared MTI with paper-based [29,31,32], one with paper-based as well as standard treatment [30], and the remaining studies compared MTI with standard treatment.

Four studies used MTI with a mobile app platform (MTI-App) and the remaining 7 used MTI with a short message service platform (MTI-SMS). Interactivity, that is, how often the person sent or received content from the MTI, ranged from daily to weekly, with daily being the most common (55%, 6/11) followed by twice daily (36%, 4/11). Seven of the studies used tailoring in the form of feedback based on individuals' symptoms [28-34]. One study sent tailored SMS based on faulty illness perceptions [35]. Four studies explicitly stated the application of behavior change theories when designing their MTI [32,35-37]. These

included the Health Belief Model [38], Illness Perceptions for Adherence [39], and Monitoring in Chronic Disease [40]. Seven studies did not report a theoretical model for their MTI [28-31,33,34,41].

Table 1. Study population characteristics of included studies.

Lead author	Year	Country	N (Duration in weeks)	Mean age in years (range)	Female %	Asthma severity
Ostojic [31]	2005	Croatia	16 (16)	25 (18+)	44 (7/16)	Moderate (?[a])
Liu [29]	2007	Taiwan	120 (24)	52 (18+)	51 (45/89)	Moderate to severe (?)
Prabhakaran [33]	2010	Singapore	120 (12)	55 (21+)	59.2 (71/120)	Poorly controlled (94%)
Strandbygaard [41]	2010	Denmark	26 (12)	32 (18-45)	46 (12/26)	Moderate to severe (69%)
Lv [30]	2012	China	150 (12)	38 (18-65)	42 (30/71)	Moderate to severe (73%)
Petrie [35]	2012	UK	147 (18)	? (16-45)	68.0 (100/147)	Nonadherence (100%)
Ryan [32]	2012	UK	288 (24)	49 (12+)	62.5 (180/288)	Poorly controlled (100%)
Yun [37]	2012	USA	30 (15)	14 (10-16)	47 (7/15)	Moderate to severe (?)
Yun [36]	2013	USA	30 (16)	13 (10-16)	57 (12/21)	Moderate to severe (?)
Cingi [28]	2015	Turkey	136 (12)	33 (25-41)	53 (47/89)	Mild to severe (?)
Zairina [34]	2016	Australia	72 (24)	31 (18+)	100 (72/72)	Moderate to severe (58%)

[a]Value could not be identified in the study.

Classification of BCT

There were 10 distinct behaviour change techniques identified within the included studies across 8 categories (Tables 2 and 3) “Instruction on how to monitor symptoms and medication” as well as “Information on health consequences” appeared to represent the standard education provided to asthma participants as part of their standard treatment and was used across all groups. All MTI-App used self-monitoring of symptoms or medication, compared with 44% (4/9) of MTI-SMS. Prompt or cues were used in 78% (7/9) of MTI-SMS compared with 25% (1/4) of MTI-App. Action planning was a lot more common in MTI-App (75%, 3/4) than MTI-SMS (11%, 1/9). Paper-based monitoring utilized similar BCT as MTI-App but did not provide feedback or social support from an external source. “Feedback” was categorized as individuals receiving feedback based on their symptoms or medication being monitored. “Social support” (practical) differed from feedback, as the participant could send a message to their physician or the investigators for responses to questions they had.

Table 2. Classification of behavior change technique taxonomy for each included study.

Lead author	Year	MTI[a]Platform	MTI	Comparator Type
Ostojic [31]	2005	MTI-SMS[b]	2.3; 2.6; 2.7; 3.2; 4.1; 5.1	Paper-based: 2.3; 4.1; 5.1
Liu [29]	2007	MTI-App[c]	1.4; 2.3; 2.6; 4.1; 5.1	Paper-based: 1.4; 2.3; 2.6; 4.1; 5.1
Prabhakaran [33]	2010	MTI-SMS	2.3; 2.7; 3.2; 4.1; 7.1	Standard treatment: 4.1
Strandbygaard [41]	2010	MTI-SMS	4.1; 5.1; 7.1	Standard treatment: 4.1; 5.1
Lv [30]	2012	MTI-SMS	1.4; 3.2; 4.1; 5.1; 7.1	Paper-based: 1.4; 2.3; 4.1; 5.1 Standard treatment: 4.1; 5.1
Petrie [35]	2012	MTI-SMS	4.2; 5.1	Standard treatment: ?
Ryan [32]	2012	MTI-App	1.4; 2.3; 2.7; 3.2; 4.1; 5.1; 6.1	Paper-based: 1.4; 2.3; 4.1; 6.1
Yun [37]	2012	MTI-SMS	2.3; 2.7; 5.1; 7.1	MTI: 2.3; 7.1 Standard treatment: ?
Yun [36]	2013	MTI-SMS	2.3; 2.7; 5.1; 7.1	MTI: 5.1; 7.1 Standard treatment: ?
Cingi [28]	2015	MTI-App	2.3; 3.2; 4.1; 7.1	Standard treatment: 4.1
Zairina [34]	2016	MTI-App	1.4; 2.3; 2.7; 3.2; 5.1	Standard treatment: 5.1

[a]MTI: Mobile technology interventions.

[b]MTI-SMS: Mobile technology interventions with short message service platform.

[c]MTI-App: Mobile technology interventions with mobile app platform.

Table 3. Behavior change technique taxonomy used by study group.

Behavior change technique taxonomy	MTI-App[a] n (%) (N=4)	MTI-SMS[b] n (%) (N=9)	MTI[c]-All n (%) (N=13)	Paper-based n (%) (N=4)	Standard treatment n (%) (N=8)
1.4 Action planning	3 (75%)	1 (11%)	4 (31%)	3 (75%)	0 (0%)
2.3 Self-monitoring of behavior	4 (100%)	4 (44%)	8 (62%)	4 (100%)	0 (0%)
2.6 Biofeedback	1 (25%)	1 (11%)	2 (15%)	1 (25%)	0 (0%)
2.7 Feedback on outcomes of behavior	2 (50%)	4 (44%)	6 (46%)	0 (0%)	0 (0%)
3.2 Social support (practical)	3 (75%)	3 (33%)	5 (46%)	0 (0%)	0 (0%)
4.1 Instruction on how to perform the behavior	3 (75%)	4 (44%)	7 (54%)	4 (100%)	4 (50%)
4.2 Information about antecedents	0 (0%)	1 (11%)	1 (8%)	0 (0%)	0 (0%)
5.1 Information about health consequences	3 (75%)	8 (89%)	11 (85%)	3 (75%)	3 (38%)
6.1 Demonstration of the behavior	1 (25%)	0 (0%)	1 (8%)	1 (25%)	0 (0%)
7.1 Prompts or cues	1 (25%)	7 (78%)	8 (62%)	0 (0%)	0 (0%)

[a]MTI-App: Mobile technology interventions with mobile app platform.

[b]MTI-SMS: Mobile technology interventions with short message service platform.

[c]MTI: Mobile technology interventions.

Risk of Bias Assessment

The results of the Cochrane Risk of Bias assessment are shown in Figures 2 and 3, with complete details presented in Multimedia Appendix 4.

Seven studies reported using appropriate random sequence generation methods (computer-generated random allocation, random order on presentation, or drawn from envelope); 4 studies did not specify the method of randomization [29,30,36,37]. Three studies reported concealment of allocation [32,34,35]. The remaining 8 studies were unclear as to whether allocation was concealed. Only 2 studies addressed blinding of participants and outcome assessors [32,34]. Although not stated, we assume that the remaining 9 studies did not carry out blinding of participants or outcome assessors.

Six of the studies had high risk of incomplete outcome data, due to varying rates of dropout or missing data between the groups [28,30,34,36], or because those who dropped out had higher medication adherence [35] or dropped out due to difficulty with MTI [29]. In one study, all participants completed the study [31], and the remaining 4 studies had low and similar dropout rates between the groups. One study provided a study protocol, which inferred low risk of selective reporting [34]. The remaining studies were unclear on potential selective reporting, but relevant variables appeared to be reported. No other potential sources of bias were identified. Overall the risk profile was quite reasonable.

Figure 2. Cochrane risk of bias assessment: risk of bias dimension for each included study.

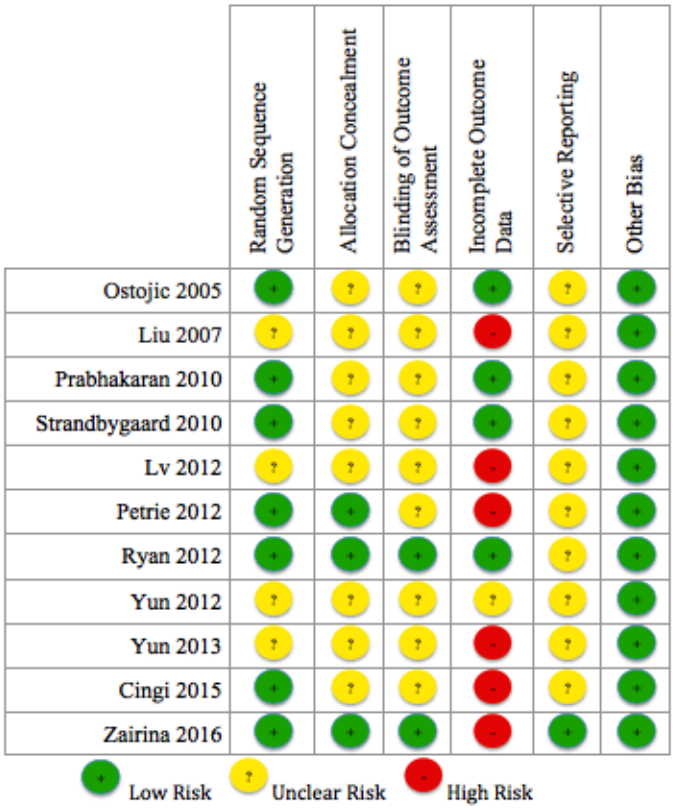
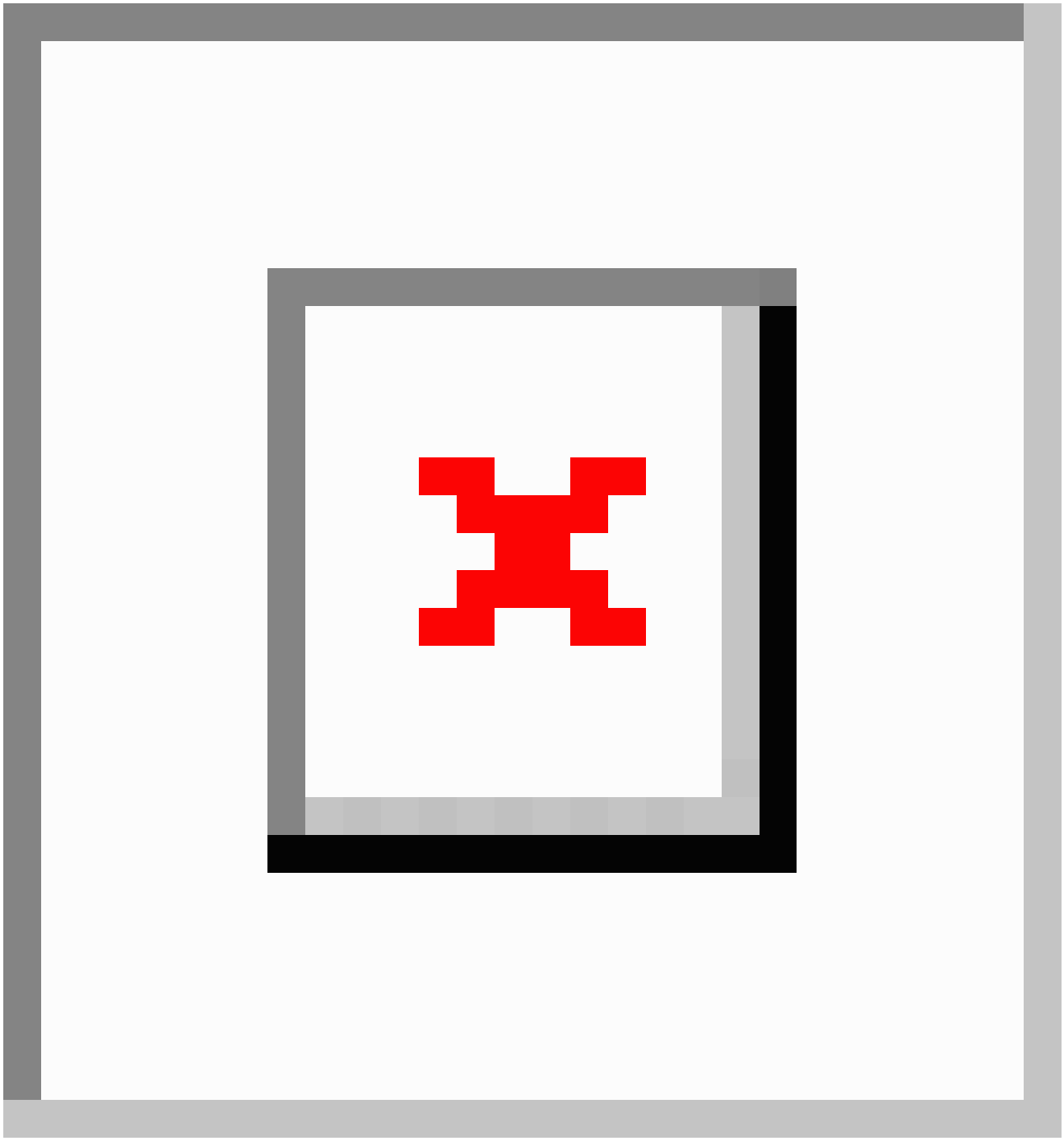


Figure 3. Cochrane risk of bias assessment: summary of each risk of bias item.



Clinical Outcomes

The studies reported adherence to medication as well as symptoms or diary entries over a time period between 12 and 24 weeks. Adherence was assessed either via prescription data or self-report throughout the intervention or at the end of the study. Four studies reported percentage adherence to preventer medication at end of study [30,32,35] or change over the study period [41]. The remaining two provided average inhaled cortico-steroid dosage at the end of study [29,31] . Two studies reported percentage adherence at end of study to symptom monitoring and action plan [29] and peak expiratory flow rate monitoring [31].

Five studies provided data on lung function in the form of mean Forced Expiratory Volume in 1 second predicted at end of the study [29,32] or change over the study period [30,34,41]. Four studies provided data on change in Quality of Life (QoL) over the study period, using the full or mini version of the Asthma-Specific Quality of Life Questionnaire [30,32,34,41]. One study provided mean scores, at the end of the study period only, using the Short-Form-12 Questionnaire—Physical Component Score [29]. Four studies provided data on change in asthma control over the study period, using the Asthma Control Questionnaire [32,34,41] and the Perceived Control of Asthma Questionnaire [30]. Three studies provided the percentage of individuals with well-controlled asthma at end

of the study [28,33,34], while 6 studies reported the percentage of unscheduled visits at end of the study [28-33].

Five studies provided the mean percentage of engagement, which ranged from 72% to 99% [29,31,33,36,37]. Three studies did not provided data on engagement with the MTI [32,34] or only median data were available [28]. A further two studies [30,41] did not have adequate engagement data as the MTI incorporated an SMS prompt that did not require engagement with the system.

MTI Effectiveness by Adherence and Health Outcome

Table 4 shows the effect of MTI on adherence based on the random effects model. Tests of heterogeneity suggested low variation among the true effects between the remaining studies. However, the width of the confidence intervals suggests an imprecise estimate of heterogeneity and caution should be taken when interpreting these results.

There was no evidence for a difference in the standardized mean medication adherence or symptom-monitoring adherence in studies comparing MTI with paper-based (Figures 4 and 5, respectively). Individuals using MTI had a significantly higher standardized mean medication adherence compared with standard treatment (Figure 6), suggesting a moderate positive effect.

Table 5 shows the effect of MTI on clinical outcomes based on the random effects model. Test of heterogeneity suggested high variation in asthma control between MTI and standard treatment studies as well as in unscheduled visits between MTI and paper-based studies. There was moderate to low variation among the true effects between the remaining studies, but the width of the confidence intervals suggests an imprecise estimate of heterogeneity and caution should be taken when interpreting these results.

There was no evidence for a difference in mean lung function, QoL, asthma control, or percentage of unscheduled visits in studies comparing MTI with paper-based monitoring (Figures 7-10). Similarly, there was no evidence for a difference in mean lung function or asthma control in studies comparing MTI with standard treatment (Figures 11 and 12, respectively). However, there was evidence for a significantly higher standardized mean QoL as well as lower mean percentage of unscheduled visits and more well-controlled asthma in studies comparing MTI with standard treatment (Figures 13-15)

Table 4. Hedges *g* and tests of heterogeneity of mobile technology intervention (MTI) for adherence.

Adherence	*k* [a]	N[b]	Hedges *g* (95%CI[c])	*P* [d]	*Q* [e]	*P* [f]	I^{2g}(95% CI)
MTI vs Paper-based							
Medication	4	450	0.16 (-0.03 to 0.34)	.10	1.08	.78	<.01 (<.01-72.42)
Symptoms	2	136	-0.11 (-0.45 to 0.22)	.51	0.22	.64	<.01 (<.01-99.55)
MTI vs Standard treatment							
Medication	3	169	0.63 (0.31 - 0.94)	<.001	0.53	.77	<.01 (<.01-89.65)

[a]Number of studies.

[b]Total sample size across included studies.

[c]95% CIs around the Hedges *g* effect size.

[d]*P* value of Hedges *g* effect size.

[e]Test of heterogeneity.

[f]*P* value of test for heterogeneity.

[g]Percentage of total variability due to heterogeneity.

Table 5. Hedges *g* and tests of heterogeneity of mobile technology intervention (MTI) for clinical outcomes.

Outcome	*k* [a]	N[b]	Hedges *g* (95%CI[c])	*P* [d]	*Q* [e]	*P* [f]	I^{2g}(95% CI)
MTI vs Paper-based							
Lung function	3	162	0.16 (-0.28 to 0.60)	.48	3.1	.21	41.63 (<.01- 97.54)
QoL[h]	3	347	0.33 (-0.08 to 0.74)	.12	6.79	.03	67.93 (<.01-99.04)
Asthma control	2	335	0.16 (−0.26 to 0.57)	.46	2.24	.14	55.28 (<.01-99.56)
Unscheduled visits	4	443	-0.49 (-1.26 to 0.27)	.21	35.5	<.001	90.51 (67.49-99.4)
MTI vs Standard Treatment							
Lung function	3	133	0.23 (-0.28 to 0.73)	.38	3.84	.15	46.15 (<.01-99.07)
QoL[h]	3	133	0.64 (0.19 - 1.08)	.01	3.29	.19	31.33 (<.01-98.88)
Asthma control	3	133	0.00 (-0.87 to 0.87)	>.99	11.79	.002	81.27 (33.23-99.50)
Well controlled	3	273	0.45 (0.20 - 0.69)	<.001	1.43	.49	<.01 (<.01-96.19)
Unscheduled visits	3	248	-0.64 (-0.90, to 0.38)	<.001	0.66	.72	<.01 (<.01-94.57)

[a]Number of studies.

[b]Total sample size across included studies.

[c]95% CIs around the hedges *g* effect size.

[d]*P* value of Hedges *g* effect size.

[e]Test of heterogeneity.

[f]*P* value of test for heterogeneity.

[g]Percentage of total variability due to heterogeneity.

[h]QoL: Quality of Life.

Figure 4. Forest plot of the standardized mean difference in medication adherence between MTI and Paper-based group (PB). MTI: mobile technology intervention.

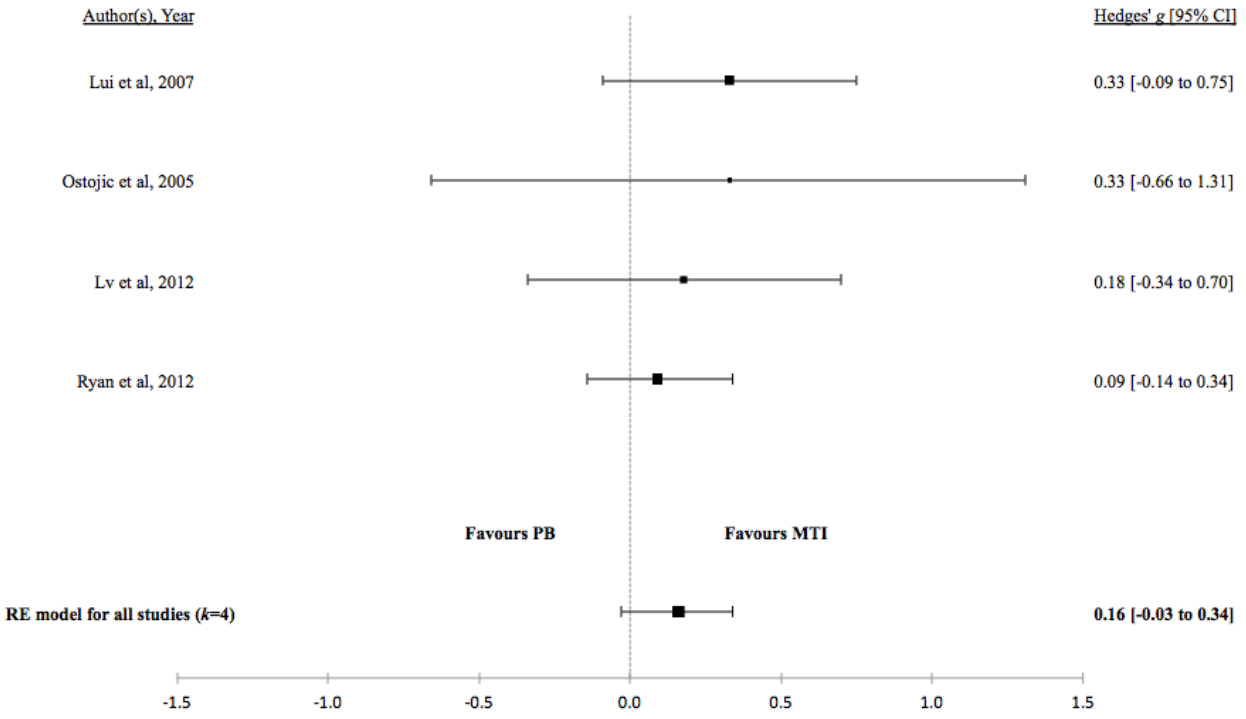


Figure 5. Forest plot of the standardized mean difference in symptom or diary adherence between MTI and paper-based group (PB). MTI: mobile technology intervention.

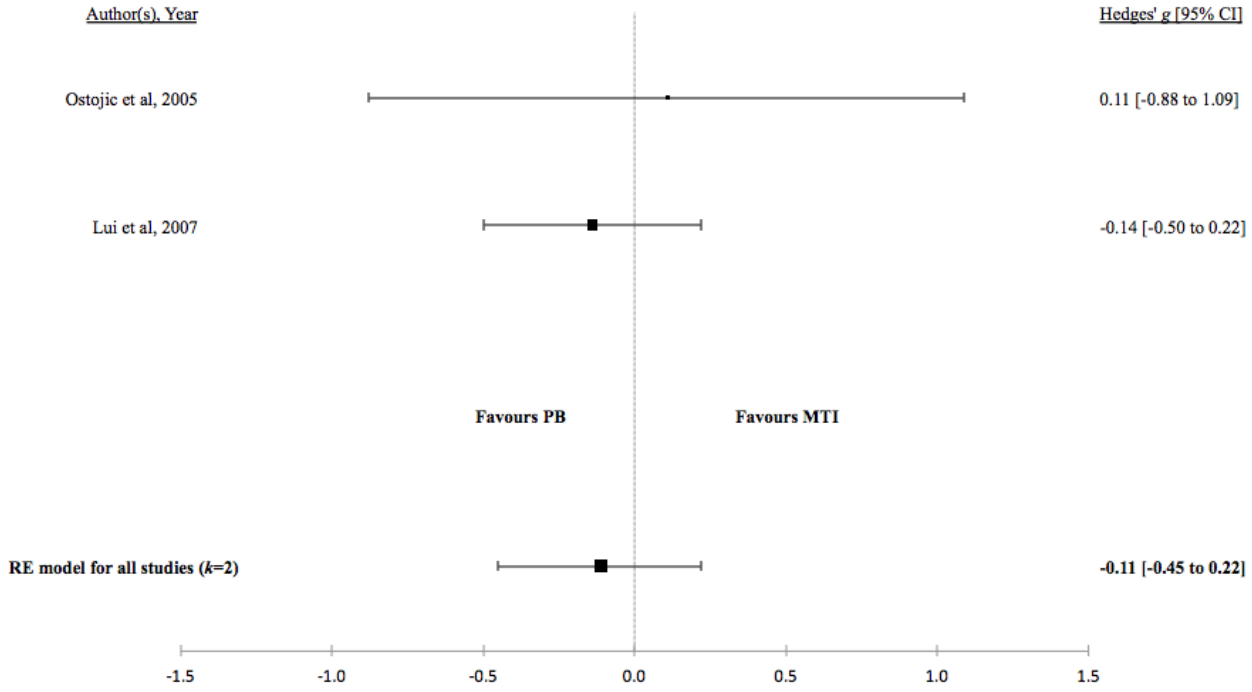


Figure 6. Forest plot of the standardized mean difference in medication adherence between MTI and standard treatment group. MTI: mobile technology intervention.

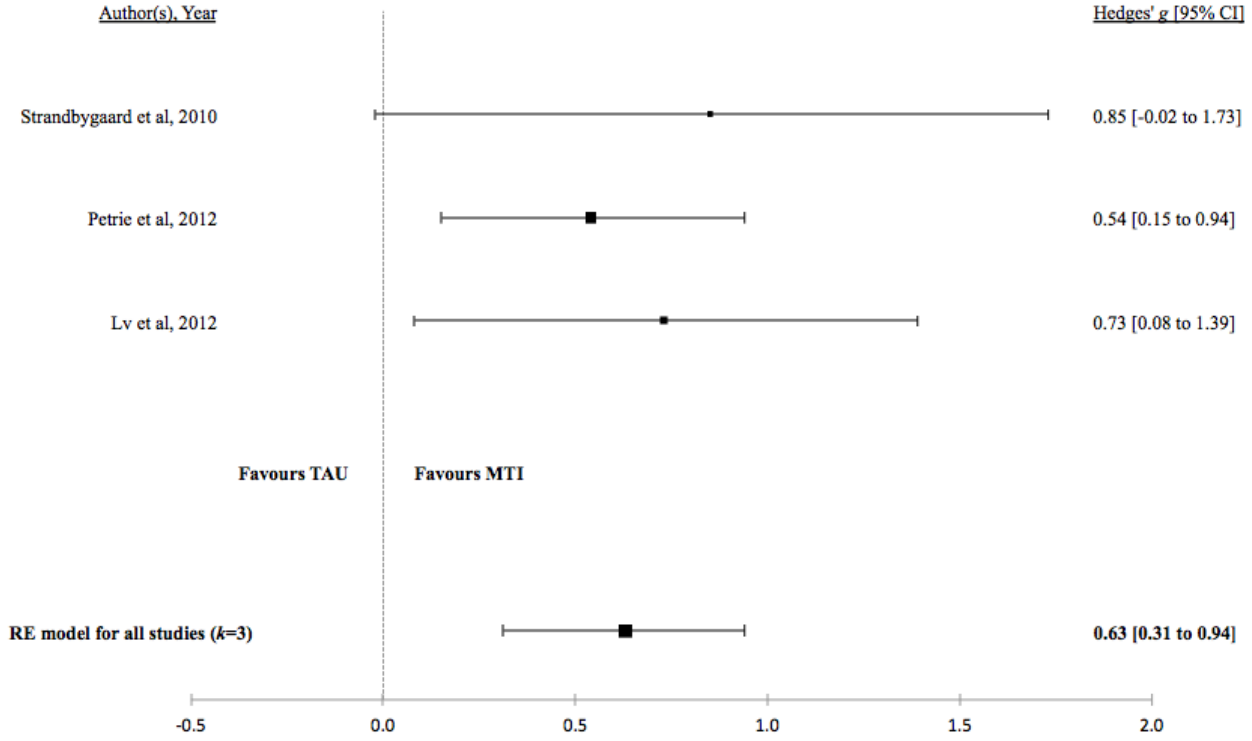


Figure 7. Forest plot of the standardized mean difference in lung function between MTI and paper-based group. MTI: mobile technology intervention.

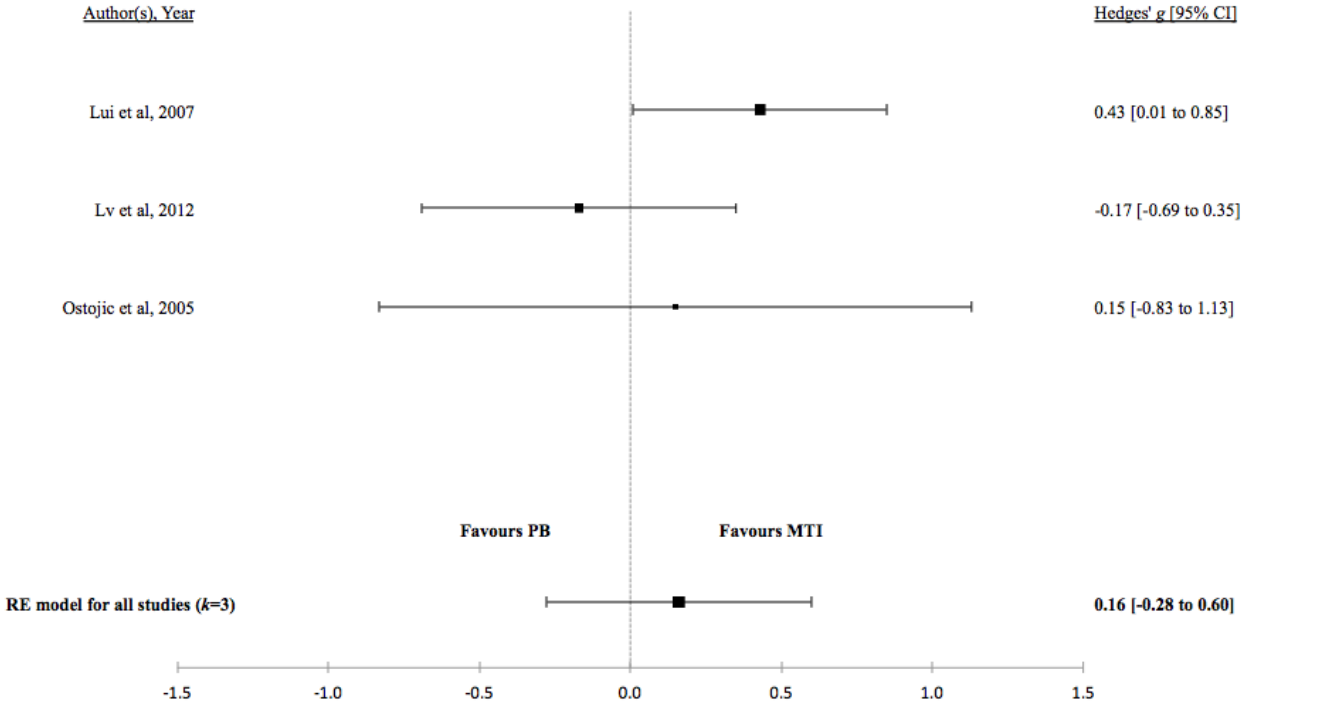


Figure 8. Forest plot of the standardized mean difference in quality of life between MTI and paper-based group (PB). MTI: mobile technology intervention.

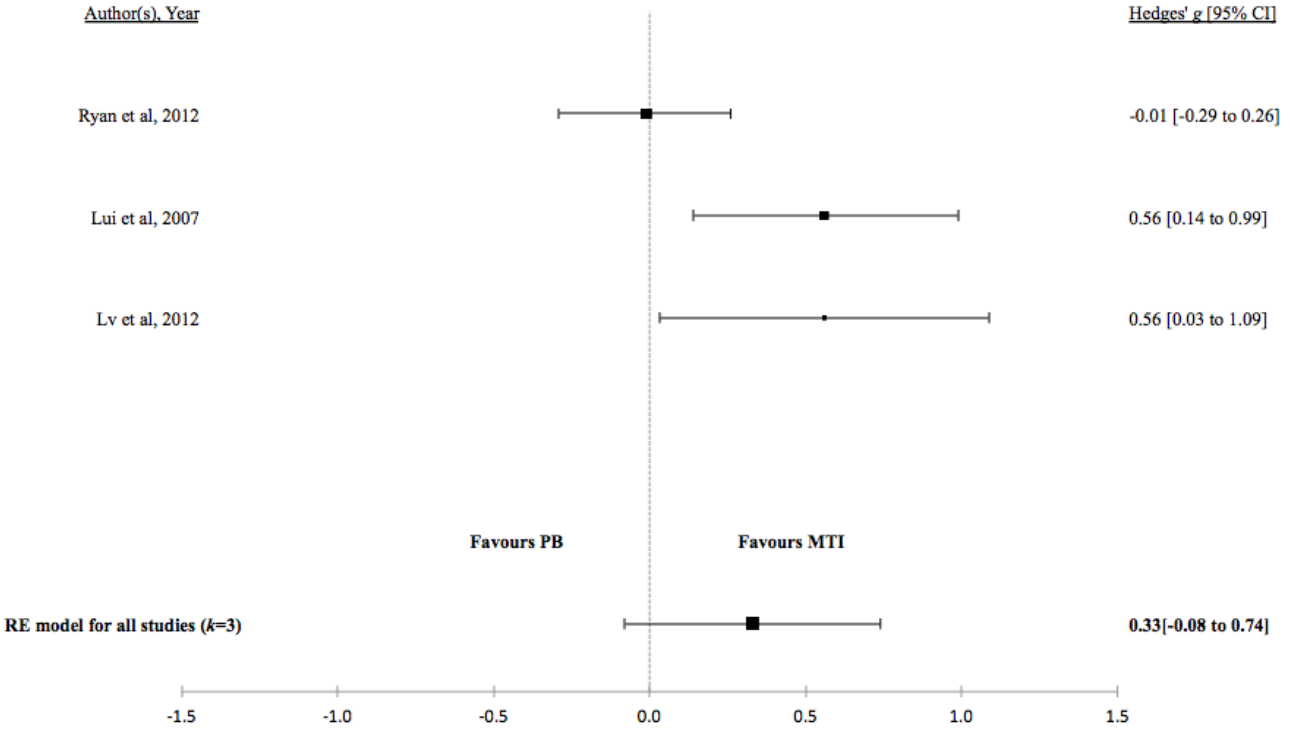


Figure 9. Forest plot of the standardized mean difference in asthma control between MTI and paper-based group (PB). MTI: mobile technology intervention.

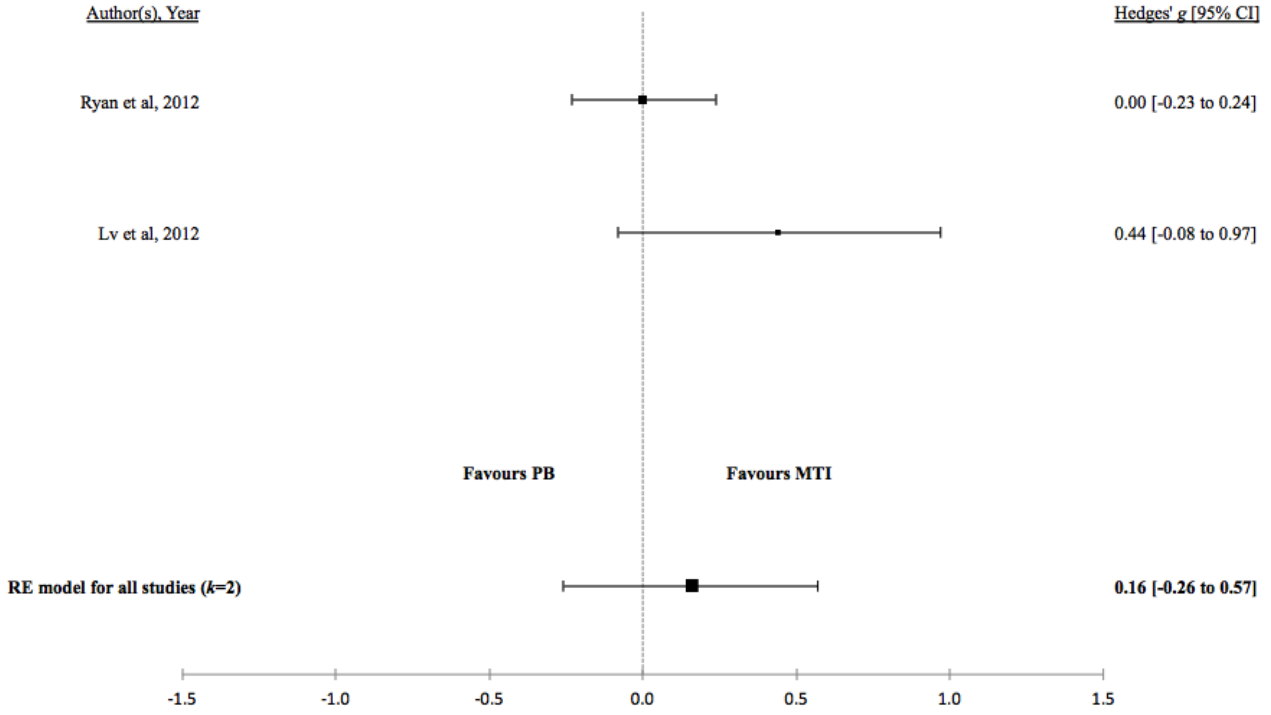


Figure 10. Forest plot of the standardized mean difference in unscheduled visits between MTI and paper-based group. MTI: mobile technology intervention.

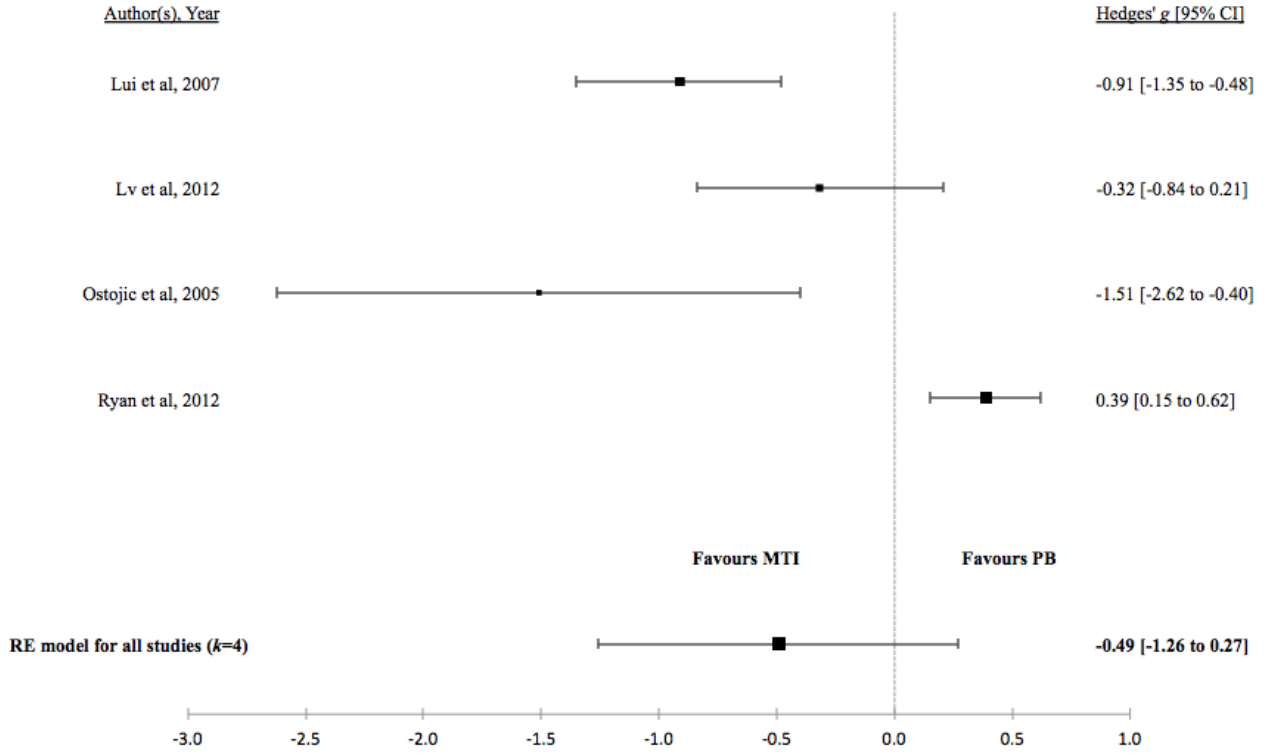


Figure 11. Forest plot of the standardized mean difference in lung function between MTI and standard treatment. MTI: mobile technology intervention.

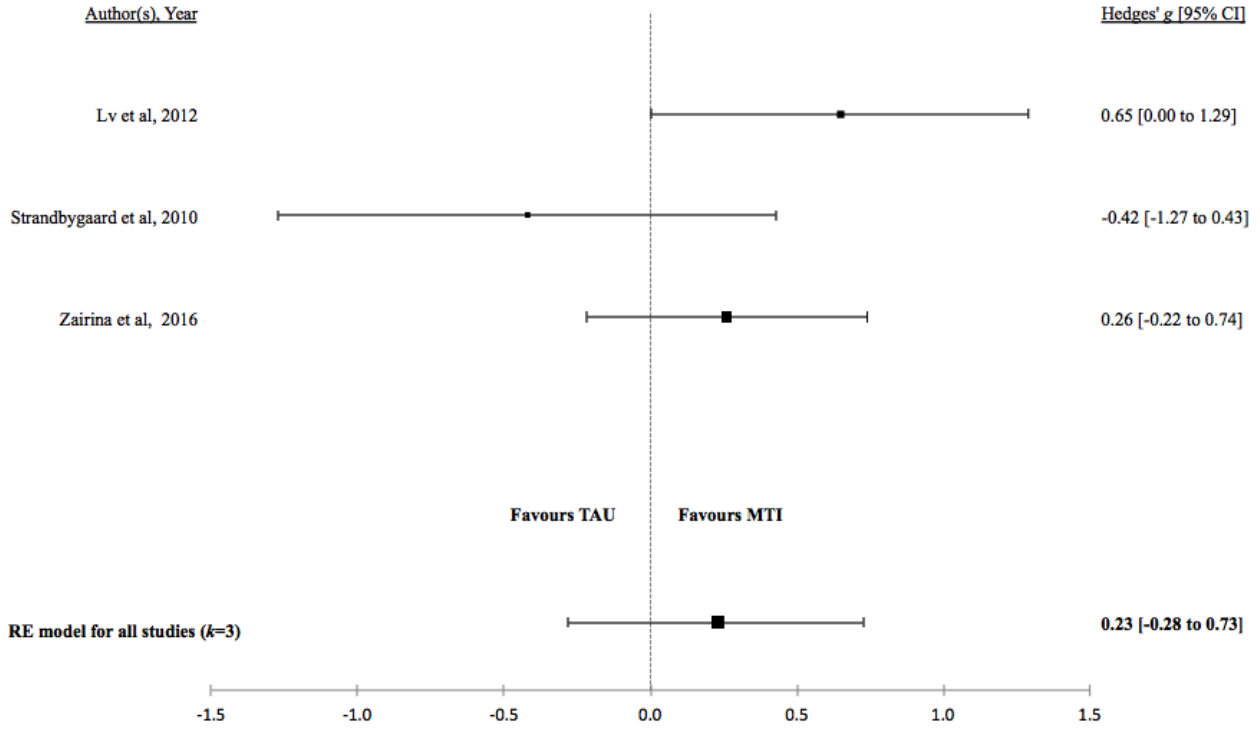


Figure 12. Forest plot of the standardized mean difference in asthma control between MTI and standard treatment. MTI: mobile technology intervention.

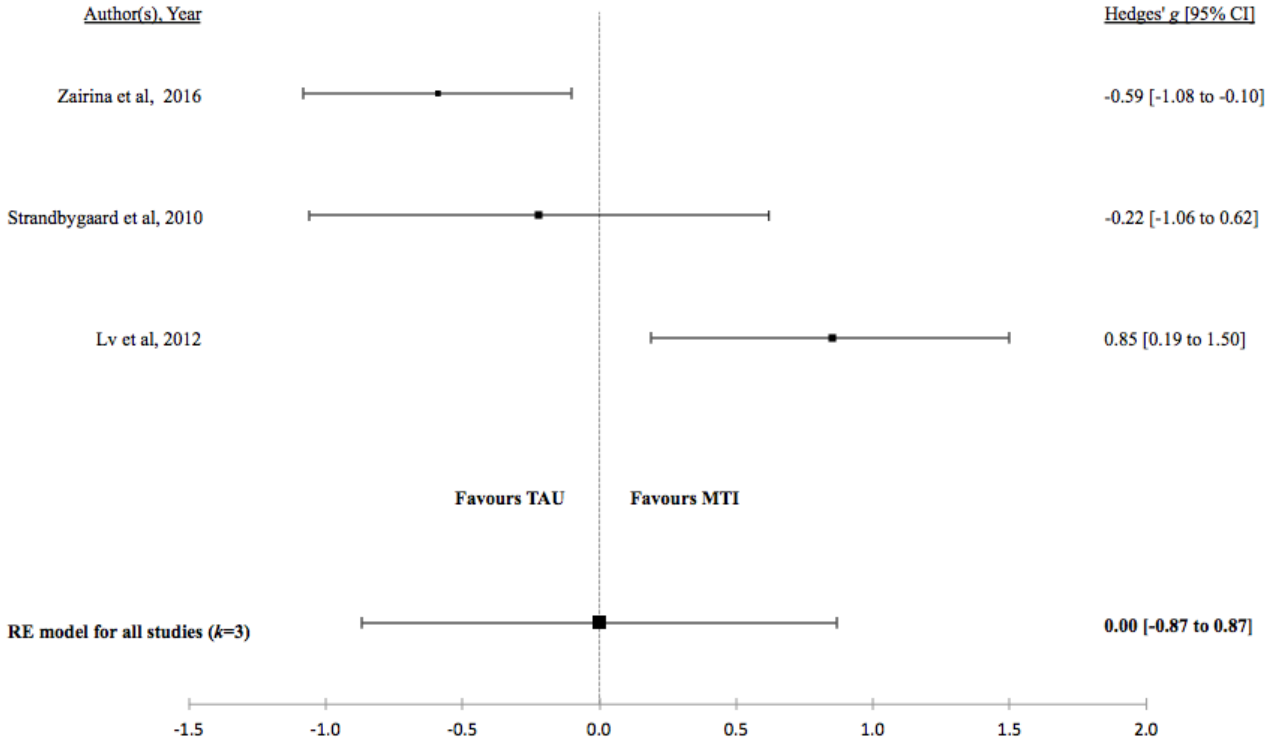


Figure 13. Forest plot of the standardized mean difference in quality of life between MTI and standard treatment. MTI: mobile technology intervention.

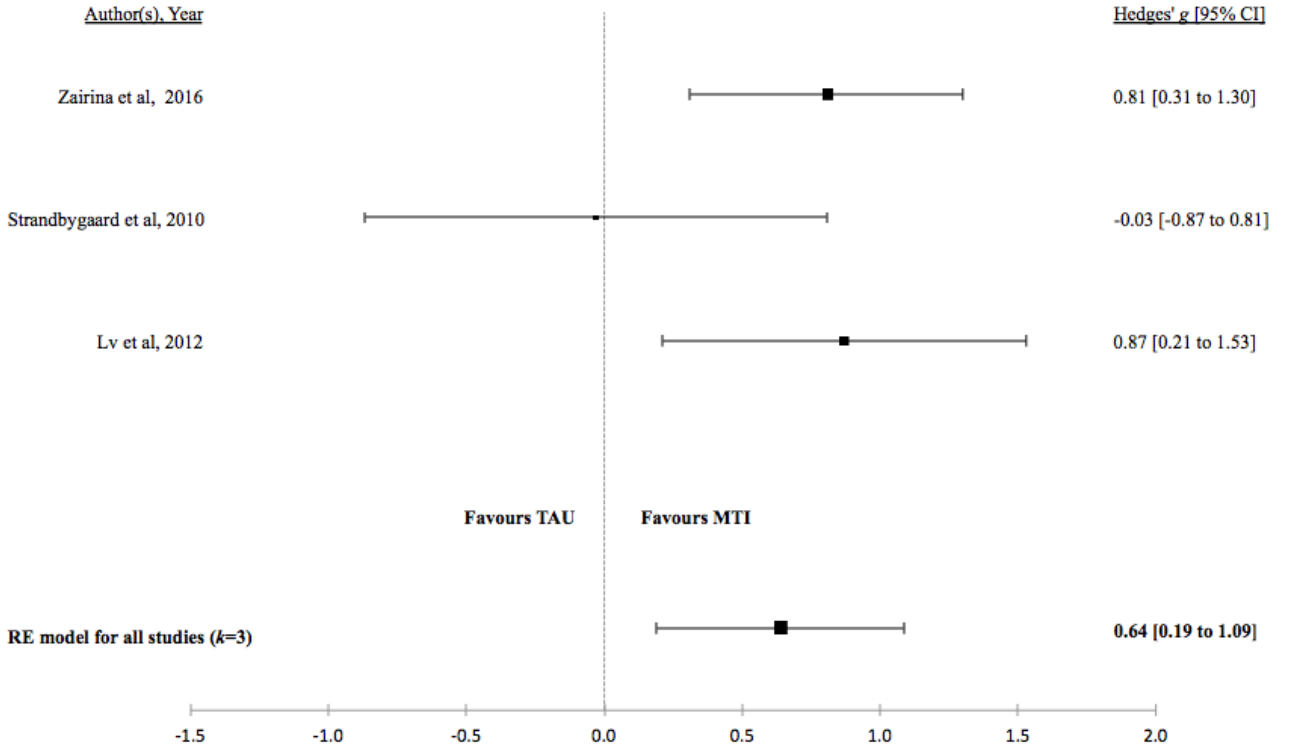


Figure 14. Forest plot of the standardized mean difference in well controlled asthma between MTI and standard treatment. MTI: mobile technology intervention.

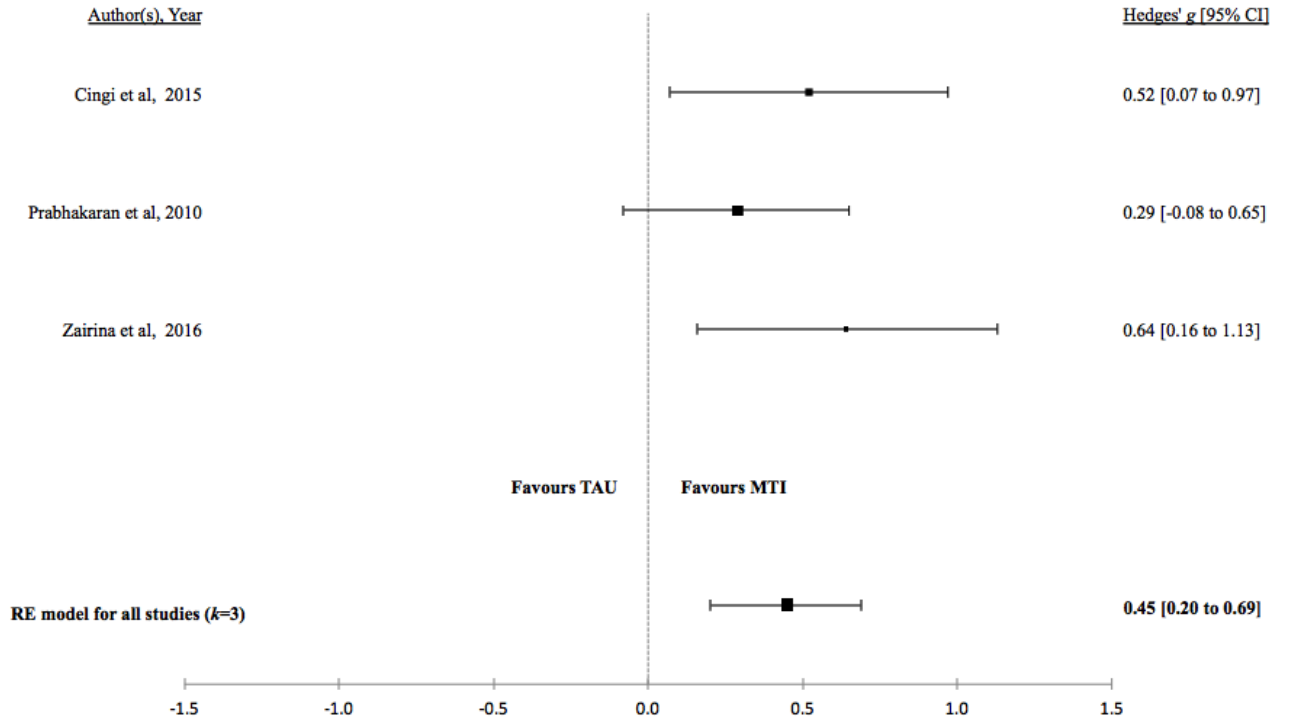
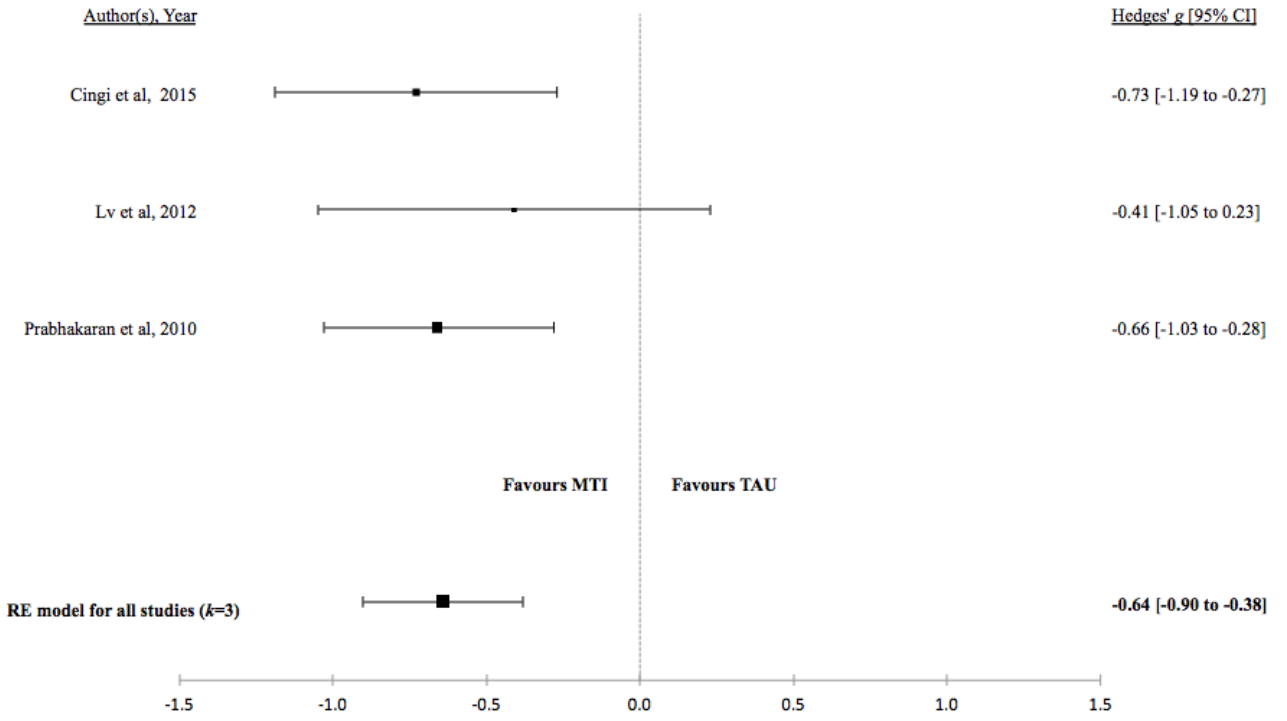


Figure 15. Forest plot of the standardized mean difference in unscheduled visits between MTI and standard treatment. MTI: mobile technology intervention.



MTI and Publication Bias: Funnel Plots

Across the relationships, symmetry of funnel plots was difficult to assess due to the small number of studies (ranging from *N*=2 -6, mean = 3). The number of studies falling outside of the funnel varied from 0 to 2 (refer to Multimedia Appendix 5 for funnel plots). Two of the 4 relationships with studies outside of the funnel (unscheduled visits between MTI and paper-based as well as asthma control between MTI and standard treatment) also had high heterogeneity, indicating that moderators may be present.

Moderation of BCT on Adherence and Clinical Outcomes

Tables 6-8 show the results of the moderator analyses. Overall, there was not adequate evidence to determine the type or number of BCT as moderators of adherence or clinical outcomes in studies comparing MTI with standard treatment. However, although the effect of number of BCT used did not reach statistical significance, it did fully explain the variation in QoL.

Table 6. Categorical moderation analysis for within group behavior change technique (BCT) type on relationship between mobile technology intervention, adherence, and clinical outcomes.

Outcome	BCT Type	*k* [a]	Hedges *g* (95%CI)	*P* [b]
Adherence	Other	1	0.54 (0.15-0.94)	.007
	Prompt Only	2	0.78 (0.25-1.30)	.004
Lung Function	Monitor + feedback	1	0.26 (−1.10-1.62)	.71
	Prompt Only	2	0.17 (−0.62 to 0.97)	.78
QoL[c]	Monitor + feedback	1	0.81 (−0.31 to 1.91)	.15
	Prompt Only	2	0.46 (−0.42 to 1.34)	.30
Asthma Control	Monitor + feedback	1	-0.59 (−1.95 to 0.77)	.40
	Prompt Only	2	0.35 (−0.7 to 1.39)	.51
Well Controlled	Monitor + feedback	2	0.43 (0.09-0.77)	.01
	Monitor Only	1	0.52 (−0.52 to 0.70)	.05

[a]Number of studies.

[b]*P* value of Hedges *g* effect size.

[c]QoL: Quality of Life.

Table 7. Categorical moderation analysis for between group behavior change technique (BCT) type on relationship between mobile technology intervention, adherence, and clinical outcomes.

Outcome	*k* [a]	BCT Type		Other moderators		I^2(95% CI)[f]
		QM [b]	*P* [c]	*QE* [d]	*P* [e]	
Adherence	3	0.48	.49	0.05	.83	<.01 (<.01-97.92)
Lung function	3	0.02	.90	3.84	.05	73.98 (<.01-99.85)
QoL[g]	3	0.23	.63	2.72	.10	63.29 (<.01-99.85)
Asthma control	3	1.14	.29	3.82	.05	73.80, (<.01-99.85)
Well controlled	3	0.08	.78	1.30	.25	23.07 (<.01-99.92)
Unscheduled visits[h]	3					

[a]Number of studies.

[b]Omnibus test of moderator (BCT Type).

[c]*P* value of test of moderator (BCT Type).

[d]Omnibus Test of other moderators not considered in model.

[e]*P* value of test for other moderators.

[f]Percentage of total variability due to heterogeneity.

[g]QoL: Quality of Life.

[h]Mixed effects model would not fit as there was only one study per BCT Type, in the analysis.

Table 8. Continuous moderation analysis for number of behavior change technique (BCT) on relationship between mobile technology intervention, adherence, and clinical outcomes.

Outcome	*k* [a]	Hedges *g* (95%CI)	Test of BCTT[b]#		Test for other moderators		I^2(95% CI)[g]
			P [c]	*R* [d]	*QE* [e]	*P* [f]	
Adherence	3	0.07 (−0.30 to 0.44)	.72	NA	0.40	.53	<.01 (<.01-99.75)
Lung Function	3	0.24 (−0.23 to 0.72)	.32	.00	2.39	.12	58.14 (<.01->99.90)
QoL[h]	3	0.26 (−0.026 to 0.59)	.11	100	0.73	.39	<0.01 (<.01-99.87)
Asthma Control	3	−0.04 (−1.01 to 0.92)	.93	.00	10.65	.001	90.61 (52.83-99.80)
Well Controlled	3	−0.09 (−0.70 to 0.52)	.78	NA	1.30	.25	23.07 (<.01-99.92)
Unscheduled Visits	3	−0.03 (−0.56 to 0.49)	.90	NA	0.64	.42	<.01 (<.01-99.85)

[a]Number of studies.

[b]BCTT: Behavior change technique taxonomy

[c]*P* value of Hedges *g* effect size.

[d]Amount of variation explained by the moderator.

[e]Omnibus Test of other moderators not considered in model.

[f]*P* value of test for other moderators.

[g]Percentage of total variability due to heterogeneity.

[h]QoL: Quality of Life.

Moderation of Engagement on Adherence and Clinical Outcomes

Only 1 study provided data on engagement for studies comparing MTI with standard treatment [33], and 2 studies comparing MTI with paper-based monitoring [29,31]. Therefore, there were insufficient data to determine if engagement was a moderator between MTI and adherence or clinical outcomes. However, there were sufficient data to test the effect of MTI on attrition using a random effects model, although caution is needed as this was a secondary analysis not included in the initial plan. For what it is worth, there was no evidence for a difference in the standardized mean attrition in studies when comparing MTI with paper-based monitoring and standard treatment (Multimedia Appendix 6).

Narrative Review

A narrative review was undertaken on the 2 studies that had insufficient data for the meta - analysis (refer to Multimedia Appendix 3 for the study characteristics). Both studies examined the effect of MTI on improving lung function and quality of life compared with standard treatment [36,37]. Consistent with findings from the meta-analysis, 1 of the 2 studies reporting on QoL found higher levels using MTI compared with standard

treatment [36]. Unlike the findings of the meta-analysis, both studies reported on lung function and found improved levels using MTI compared with standard treatment [36,37].

Discussion

Principal Findings

The aim of this systematic review and meta-analysis was to determine the efficacy of MTI for asthma self-management. The findings support the effects of MTI for improving adherence and clinical outcomes in asthma [15,21] if compared with standard treatment. Furthermore, these results are consistent with the relationship between MTI, adherence, and clinical outcomes in chronic illness [15] and provide an evidence base for the content of MTI. However, more research is required to determine the role of BCT and engagement in MTI.

The current review found that MTI had a moderate effect on improving medication adherence when compared with standard treatment. Further, the current review suggests that MTI indeed do have a large effect on improving QoL and a moderate effect on individuals achieving well-controlled asthma and fewer unscheduled visits. This is in contrast to previous findings by Tran and colleagues [21] that MTI does not improve clinical outcomes. Consistent with their findings, this review found no evidence of a difference in mean lung function. In addition, the findings in this review are consistent with previous research on chronic illnesses in general that found positive changes in clinical outcomes when using MTI [15], as well as a positive association between medication adherence and improved clinical outcomes for individuals with asthma [5].

Although study duration was identified as a potential limitation to finding change in clinical outcomes [21], the current review noted similar duration in all studies, of around 3 months. The current findings are an advance on previous attempts at systematic reviews of MTI, as it features increased power due to more included studies and by excluding interventions based on telephone-call monitoring, thereby showing clearer effects of MTI. The findings also support claims by physicians working with individuals with asthma who proposed that MTI would improve the quality of care [14]. However, longer duration studies would certainly be valuable. Clinically relevant improvements in QoL, asthma control, and lung function between 3 and 12 months have been found when using Internet-based self-management of asthma compared with standard treatment over 12 months [42], which suggests extending the duration of MTI studies. Furthermore, improvements in QoL and asthma control from Internet-based self-management have been sustained for up to 1.5 years after support has ceased [43].

However, a different picture emerges when MTI effects are compared with paper-based monitoring. Here, the current review found no evidence of a difference in medication or symptom or diary adherence or clinical outcomes.

This suggests 2 indirect conclusions: First, given MTI resulted in improved adherence and clinical outcomes over standard treatment, it suggests that providing instruction on behavior and information about health consequences in itself is not sufficient to improve adherence and consequent health benefits in individuals with asthma. This is consistent with previous findings of a similar improvement in asthma knowledge and inhalation technique between Internet-based self-management and standard treatment over 12 months, concluding that baseline meetings can trigger improvements in these areas without the need for further education [42]. Second, given that there was no difference between MTI and paper-based monitoring, it could suggest improvements in medication adherence and clinical outcomes can be made by providing some feature to monitor medication or symptoms in either electronic or conventional paper form.

As in previous reviews [18], only few of the included studies reported engagement, and when reported at all was widely variable. There was insufficient detail to undertake a formal analysis of engagement, and a secondary analysis looking at individual dropout from the study found no evidence of a difference between MTI and paper-based or standard treatment. However, as MTI are probably an acceptable medium, the fact that they are more portable and potentially more accurate than a paper diary [11-13] points to their potential to help people manage and cope with their asthma.

Limitations and Future Research

Similar to previous reviews supporting the efficacy of MTI in supporting self-management of chronic illnesses in general [15], this review provides strong evidence for at least short-term efficacy of MTI for asthma management. This supports the need for longer duration studies to identify not just efficacy, but also cost effectiveness, client and physician safety, and the specific constituent behaviour change techniques within the mobile technology intervention that really matters most. Consistent with previous research [15], there was much variation in the study designs and intervention characteristics in this review, making it difficult to draw clear associations between the individual effects of the MTI, the specific study design, and included BCT.

Concerning study quality, small sample size and poor recruitment appear to be a consistent issue in research on MTI and may lead to a lack of generalizability of findings to the broader asthma population [15]. The asthma population varies in regard to asthma severity, health literacy, ability for self-management, and preference for mobile management [44], which makes generalizing findings from single studies and even reviews difficult. The exclusion of individuals with comorbidities also needs addressing as these are frequent in older individuals [1], and this may be contributing to the exclusion of older individuals within MTI research. Furthermore, as recent research [45] has shown that up to one-third of people with physician-diagnosed asthma in the past 5 years do not have current asthma; future treatment evaluations need to establish current asthma objectively, since treatment is unlikely to be effective in people without diagnosed current asthma. For this study, this means that we cannot rule out that people without diagnosed current asthma received treatment which in turn showed no effects.

In addition, this review is limited by the degree to which the included studies reported a theoretical basis for their intervention

and subsequently based their interventions on this theory. Internet and mobile-based interventions for promoting health-behavior change have been found to be more efficacious when incorporating more extensive use of theory, BCT, and using text messages as mode of delivery [46]. However, current MTI often lack a theoretical basis underlying the MTI content [15]. In this review, less than half of the included studies provided a theoretical basis for their MTI, and no study specified the BCT used. Thus, the classification of BCT according to a current taxonomy [16] is based only on the information provided in the studies, and we cannot confirm the authors' original intention.

Two studies in this review found evidence to support changes in medication adherence alongside changes in the individual's perception of asthma [35,41]. Illness perceptions play a role in asthma self-management [9], with perceived changes to beliefs around duration, control, and severity of asthma associated with the use of MTI with self-monitoring [11,13]. If MTI with self-monitoring indirectly changes asthma perceptions then this validates the Common Sense Model of Illness [10] as a theoretical framework for asthma MTI. Further exploration is warranted on whether targeting illness perceptions alone is sufficient for lasting behavior change and highlights the important related question of defining whom to target with which behaviour change technique. One of the included studies found targeting faulty illness perceptions through MTI increased medication adherence [35]. Paradoxically, this study also found those with higher adherence were less likely to remain in the study [35], suggesting that only people with partially or uncontrolled asthma due to faulty illness perceptions would benefit from this type of MTI.

Another study found individuals with milder asthma were more likely to withdraw from paper-based monitoring than standard treatment or MTI [30]. This may suggest that MTI with monitoring is more attractive than paper-based in those with milder severity. Given the current guidelines recommending action plans for asthma management, an MTI that can store and activate action plans would be suitable for any level of severity, but presumably those at most need at the more severe end of the range. Action plans are most effective at improving clinical outcomes when two to four action points are included based on asthma symptoms or peak expiratory flows [4]. The addition of inhalation devices to electronic monitoring devices, perhaps via Bluetooth, has been proposed to record inhalation use data as well as inhalation technique [44]. However, increasing the complexity of the mobile app will also lead to increases in cost and risk to the individual with asthma and the physician, if not adequately tested [44,47].

Perhaps the main limitation of this review was the small number of included studies, which meant that the power of the tests was relatively low, with a high chance of Type-2 statistical errors [24]. On the other hand, the fact that differences were found at all indicates stable effects. Tests of heterogeneity and effect sizes both produced large confidence intervals, again reflecting imprecise estimates. Furthermore, tests of publication bias could not be performed. To validate the findings of this review and extend the ability to do robust analyses before public use, longer and larger studies are required that incorporate a theoretical basis and behavior change techniques [46], and that follow guidelines for reporting MTI [48].

Implications and Conclusion

MTI for asthma management can improve medication adherence and quality of life, decrease unscheduled visits, and increase the likelihood of achieving well-controlled asthma compared with standard treatment alone. In addition, MTI appear to be equally as efficacious as paper-based monitoring at achieving higher medication adherence and clinical outcomes. Better reporting of BCT and further research into long-term efficacy of MTI for adherence behavior and clinical outcomes is needed to create an evidence base for specific behaviour change techniques that best support individuals with asthma.

Finally, implementation of mobile technology interventions poses a challenge: Physicians' concerns regarding increased time and resource demands, accuracy of information, physician liability, and patient confidentiality [14] must be addressed. Furthermore, as MTI are being used to make critical decisions, they must be tested for accuracy and reliability to reduce harm to the individual and potential liability for the physician [47]. Increased involvement of the physician in the development and testing phases of the MTI as well as assessment of privacy issues, and improvements in regulation and safety checks by regulatory bodies and the development of a risk assessment framework for medical mobile app is warranted [44,47].

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 15KB - mhealth_v5i5e57_app1.pdf]

Multimedia Appendix 2

Studies excluded during full-text review.

[PDF File (Adobe PDF File), 62KB - mhealth_v5i5e57_app2.pdf]

Multimedia Appendix 3

Extracted characteristics of studies in meta-analysis and narrative review.

[PDF File (Adobe PDF File), 159KB - mhealth_v5i5e57_app3.pdf]

Multimedia Appendix 4

Risk of bias analysis.

[PDF File (Adobe PDF File), 24KB - mhealth_v5i5e57_app4.pdf]

Multimedia Appendix 5

Funnel plots.

[PDF File (Adobe PDF File), 273KB - mhealth_v5i5e57_app5.pdf]

Multimedia Appendix 6

Attrition analysis.

[PDF File (Adobe PDF File), 146KB - mhealth_v5i5e57_app6.pdf]

References

1. Bateman ED, Hurd SS, Barnes PJ, Bousquet J, Drazen JM, FitzGerald M, et al. Global strategy for asthma management and prevention: GINA executive summary. Eur Respir J 2008 Jan;31(1):143-178 [FREE Full text] [doi: 10.1183/09031936.00138707] [Medline: 18166595]
2. Australian Bureau of Statistics.: Australian Bureau of Statistics; 2015. 4364.0.55.001 - National Health Survey: First Results, 2014-15 URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/4364.0.55.001 [accessed 2016-12-14] [WebCite Cache ID 6mlRDzsUp]
3. Australian Asthma Foundation.: Australian Asthma Foundation; 2015 Apr. Australian Asthma Handbook, Version 1.2 URL: https://www.asthmahandbook.org.au/about/updates/version1_2 [accessed 2016-12-14] [WebCite Cache ID 6mlRULO9j]
4. Gibson PG, Powell H. Written action plans for asthma: an evidence-based review of the key components. Thorax 2004 Feb;59(2):94-99 [FREE Full text] [Medline: 14760143]
5. Lindsay JT, Heaney LG. Nonadherence in difficult asthma - facts, myths, and a time to act. Patient Prefer Adherence 2013 Apr;7:329-336 [FREE Full text] [doi: 10.2147/PPA.S38208] [Medline: 23723690]
6. Reddel H, Sawyer S, Flood P, Everett P, Peters M. Patterns of asthma control and inhaled corticosteroid (ICS) use in Australians living with asthma. Respirology 2014 Mar 24;19(2):77 [FREE Full text] [doi: 10.1111/resp.12263_5]
7. Correll PK, Poulos LM, Ampon R, Reddel HK, Marks GB. Australian Institute of Health and Welfare. 2015 May 05. Respiratory medication use in Australia 2003-2013: treatment of asthma and COPD URL: http://www.aihw.gov.au/publication-detail/?id=60129550778 [accessed 2016-12-15] [WebCite Cache ID 6mlRhC3Ll]
8. Kandane-Rathnayake RK, Matheson MC, Simpson JA, Tang ML, Johns DP, Mészáros D, et al. Adherence to asthma management guidelines by middle-aged adults with current asthma. Thorax 2009 Dec;64(12):1025-1031. [doi: 10.1136/thx.2009.118430] [Medline: 19703827]
9. Horne R, Weinman J. Self-Regulation and self-management in asthma: exploring the role of illness perceptions and treatment beliefs in explaining non-adherence to preventer medication. Health Psychol 2002;17(1):17-32. [doi: 10.1080/08870440290001502]
10. Leventhal H, Meyer D, Nerenz D. The common sense representation of illness danger. In: Rachman S, editor. Medical Psychology, Volume II. New York: Pergamon; 1980.
11. Cleland J, Caldow J, Ryan D. A qualitative study of the attitudes of patients and staff to the use of mobile phone technology for recording and gathering asthma data. J Telemed Telecare 2007;13(2):85-89. [doi: 10.1258/135763307780096230] [Medline: 17359572]
12. Fonseca JA, Costa-Pereira A, Delgado L, Fernandes L, Castel-Branco MG. Asthma patients are willing to use mobile and web technologies to support self-management. Allergy 2006 Mar;61(3):389-390 [FREE Full text] [doi: 10.1111/j.1398-9995.2006.01016.x] [Medline: 16436151]
13. Anhøj J, Møldrup C. Feasibility of collecting diary data from asthma patients through mobile phones and SMS (short message service): response rate analysis and focus group evaluation from a pilot study. J Med Internet Res 2004 Dec 2;6(4):e42 [FREE Full text] [doi: 10.2196/jmir.6.4.e42] [Medline: 15631966]
14. Martinasek MP, Panzera AD, Schneider T, Lindenberger JH, Bryant CA, McDermott RJ, et al. Benefits and barriers of pediatric healthcare providers toward using social media in asthma care. Am J Health Behav 2011;42(4):213-221. [doi: 10.1080/19325037.2011.10599190]

15. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
16. 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]
17. Morrissey EC, Corbett TK, Walsh JC, Molloy GJ. Behavior change techniques in apps for medication adherence: a content analysis. Am J Prev Med 2016 May;50(5):e143-e146. [doi: 10.1016/j.amepre.2015.09.034] [Medline: 26597504]
18. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
19. 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;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
20. Marcano Belisario JS, Huckvale K, Greenfield G, Car J, Gunn LH. Smartphone and tablet self management apps for asthma. Cochrane Database Syst Rev 2013;11:CD010013. [doi: 10.1002/14651858.CD010013.pub2] [Medline: 24282112]
21. Tran N, Coffman JM, Sumino K, Cabana MD. Patient reminder systems and asthma medication adherence: a systematic review. J Asthma 2014 Jun;51(5):536-543. [doi: 10.3109/02770903.2014.888572] [Medline: 24506699]
22. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
23. Malhotra S, Musgrave SD, Pinnock H, Price D, Ryan DP. The challenge of recruiting in primary care for a trial of telemonitoring in asthma: an observational study. Pragmat Obs Res 2012;3:51-55 [FREE Full text] [doi: 10.2147/POR.S34380] [Medline: 27774017]
24. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. 2011. URL: http://handbook.cochrane.org/ [accessed 2017-04-18] [WebCite Cache ID 6poGU52Dr]
25. Durlak JA. How to select, calculate, and interpret effect sizes. J Pediatr Psychol 2009 Oct;34(9):917-928 [FREE Full text] [doi: 10.1093/jpepsy/jsp004] [Medline: 19223279]
26. R Core Team. R-project. Vienna, Austria: R Foundation for Statistical Computing; 2011. R: A language and environment for statistical computing URL: http://www.r-project.org/ [accessed 2017-04-08] [WebCite Cache ID 6paCIvDLF]
27. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychological Methods 1998;3(4):486-504. [doi: 10.1037//1082-989X.3.4.486]
28. Cingi C, Yorgancioglu A, Cingi CC, Oguzulgen K, Muluk NB, Ulusoy S, et al. The "physician on call patient engagement trial" (POPET): measuring the impact of a mobile patient engagement application on health outcomes and quality of life in allergic rhinitis and asthma patients. Int Forum Allergy Rhinol 2015 Jun;5(6):487-497. [doi: 10.1002/alr.21468] [Medline: 25856270]
29. Liu W, Huang C, Wang C, Lee K, Lin S, Kuo H. A mobile telephone-based interactive self-care system improves asthma control. Eur Respir J 2011 Feb;37(2):310-317 [FREE Full text] [doi: 10.1183/09031936.00000810] [Medline: 20562122]
30. Lv Y, Zhao H, Liang Z, Dong H, Liu L, Zhang D, et al. A mobile phone short message service improves perceived control of asthma: a randomized controlled trial. Telemed J E Health 2012;18(6):420-426. [doi: 10.1089/tmj.2011.0218] [Medline: 22667695]
31. Ostojic V, Cvoriscec B, Ostojic SB, Reznikoff D, Stipic-Markovic A, Tudjman Z. Improving asthma control through telemedicine: a study of short-message service. Telemed J E Health 2005 Feb;11(1):28-35. [doi: 10.1089/tmj.2005.11.28] [Medline: 15785218]
32. Ryan D, Price D, Musgrave SD, Malhotra S, Lee AJ, Ayansina D, et al. Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: multicentre randomised controlled trial. BMJ 2012;344:e1756 [FREE Full text] [Medline: 22446569]
33. Prabhakaran L, Chee WY, Chua KC, Abisheganaden J, Wong WM. The use of text messaging to improve asthma control: a pilot study using the mobile phone short messaging service (SMS). J Telemed Telecare 2010;16(5):286-290. [doi: 10.1258/jtt.2010.090809] [Medline: 20576744]
34. Zairina E, Abramson MJ, McDonald CF, Li J, Dharmasiri T, Stewart K, et al. Telehealth to improve asthma control in pregnancy: a randomized controlled trial. Respirology 2016 Jul;21(5):867-874. [doi: 10.1111/resp.12773] [Medline: 27037722]
35. Petrie KJ, Perry K, Broadbent E, Weinman J. A text message programme designed to modify patients' illness and treatment beliefs improves self-reported adherence to asthma preventer medication. Br J Health Psychol 2012 Feb;17(1):74-84. [doi: 10.1111/j.2044-8287.2011.02033.x] [Medline: 22107110]
36. Yun T, Arriaga R. A text message a day keeps the pulmonologist away. 2013 Presented at: CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2013; Paris, France p. 1769-1778 URL: http://dl.acm.org/citation.cfm?id=2466233 [doi: 10.1145/2470654.2466233]
37. Yun T, Joeng H, Hill T, Lesnick B, Brown R, Abowd G. Using SMS to provide continuous assessment and improve health outcomes for children with asthma. 2012 Presented at: IHI '12 Proceedings of the 2nd ACM SIGHIT International Health

Informatics Symposium; January 28-30, 2012; Miami, FL p. 621-630 URL: http://dl.acm.org/citation.cfm?id=2110432 [doi: 10.1145/2110363.2110432]
38. Rosenstock IM. Historical origins of the health belief model. Health Educ Behav 1974 Dec 01;2(4):328-335. [doi: 10.1177/109019817400200403]
39. Petrie KJ, Weinman J. Why illness perceptions matter. Clin Med (Lond) 2006;6(6):536-539. [Medline: 17228551]
40. Glasziou P, Irwig L, Mant D. Monitoring in chronic disease: a rational approach. BMJ 2005 Mar 19;330(7492):644-648 [FREE Full text] [doi: 10.1136/bmj.330.7492.644] [Medline: 15774996]
41. 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]
42. van der Meer V, Bakker MJ, van den Hout WB, Rabe KF, Sterk PJ, Kievit J, SMASHING (Self-Management in Asthma Supported by Hospitals, ICT, NursesGeneral Practitioners) Study Group. Internet-based self-management plus education compared with usual care in asthma: a randomized trial. Ann Intern Med 2009 Jul 21;151(2):110-120. [Medline: 19620163]
43. van Gaalen JL, Beerthuizen T, van der Meer V, van Reisen P, Redelijkheid GW, Snoeck-Stroband JB, et al. Long-term outcomes of internet-based self-management support in adults with asthma: randomized controlled trial. J Med Internet Res 2013;15(9):e188 [FREE Full text] [doi: 10.2196/jmir.2640] [Medline: 24028826]
44. van Boven JF, Trappenburg JC, van der Molen T, Chavannes NH. Towards tailored and targeted adherence assessment to optimise asthma management. NPJ Prim Care Respir Med 2015 Jul 16;25:15046 [FREE Full text] [doi: 10.1038/npjpcrm.2015.46] [Medline: 26181850]
45. Aaron SD, Vandemheen KL, FitzGerald JM, Ainslie M, Gupta S, Lemière C, Canadian Respiratory Research Network. Reevaluation of diagnosis in adults with physician-diagnosed asthma. JAMA 2017 Dec 17;317(3):269-279. [doi: 10.1001/jama.2016.19627] [Medline: 28114551]
46. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
47. Lewis TL, Wyatt JC. mHealth and mobile medical apps: a framework to assess risk and promote safer use. J Med Internet Res 2014 Sep 15;16(9):e210 [FREE Full text] [doi: 10.2196/jmir.3133] [Medline: 25223398]
48. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016;352:i1174. [Medline: 26988021]

Abbreviations

BCT: Behavior change technique
BCTT: Behavior change technique taxonomy
MTI: mobile technology intervention
MTI-App: MTI with mobile app platform
MTI-SMS: MTI with short messaging service platform
QoL: Quality of Life

Edited by G Eysenbach; submitted 14.12.16; peer-reviewed by D Ryan; comments to author 29.01.17; revised version received 01.03.17; accepted 17.03.17; published 02.05.17.

Please cite as:
Miller L, Schüz B, Walters J, Walters EH
Mobile Technology Interventions for Asthma Self-Management: Systematic Review and Meta-Analysis
JMIR Mhealth Uhealth 2017;5(5):e57
URL: https://mhealth.jmir.org/2017/5/e57/
doi:10.2196/mhealth.7168
PMID:28465281

Original Paper

Effect of a Counseling Session Bolstered by Text Messaging on Self-Selected Health Behaviors in College Students: A Preliminary Randomized Controlled Trial

Janice Sandrick[1], RD, LDN, PhD; Doreen Tracy[1], RD, LDN, MS; Arn Eliasson[2], MD; Ashley Roth[1], RD, LDN, MS; Jeffrey Bartel[3], PhD; Melanie Simko[1], BS, RD, LDN; Tracy Bowman[1], RD, LDN, MS; Karen Harouse-Bell[1], RD, LDN, MS; Mariam Kashani[2], DNP, CRNP; Marina Vernalis[2], DO

[1]Coordinated Program in Nutrition and Dietetics, Seton Hill University, Greensburg, PA, United States

[2]Integrative Cardiac Health Project, Department of Medicine, Walter Reed Military Medical Center, Bethesda, MD, United States

[3]Psychology Program, Seton Hill University, Greensburg, PA, United States

Corresponding Author:
Arn Eliasson, MD
Integrative Cardiac Health Project
Department of Medicine
Walter Reed Military Medical Center
Bldg 17A
8901 Wisconsin Avenue
Bethesda, MD, 20889
United States
Phone: 1 301 400 1111
Fax: 1 301 400 1620
Email: aheliasson@aol.com

Abstract

Background: The college experience is often the first time when young adults live independently and make their own lifestyle choices. These choices affect dietary behaviors, exercise habits, techniques to deal with stress, and decisions on sleep time, all of which direct the trajectory of future health. There is a need for effective strategies that will encourage healthy lifestyle choices in young adults attending college.

Objective: This preliminary randomized controlled trial tested the effect of coaching and text messages (short message service, SMS) on self-selected health behaviors in the domains of diet, exercise, stress, and sleep. A second analysis measured the ripple effect of the intervention on health behaviors not specifically selected as a goal by participants.

Methods: Full-time students aged 18-30 years were recruited by word of mouth and campuswide advertisements (flyers, posters, mailings, university website) at a small university in western Pennsylvania from January to May 2015. Exclusions included pregnancy, eating disorders, chronic medical diagnoses, and prescription medications other than birth control. Of 60 participants, 30 were randomized to receive a single face-to-face meeting with a health coach to review results of behavioral questionnaires and to set a health behavior goal for the 8-week study period. The face-to-face meeting was followed by SMS text messages designed to encourage achievement of the behavioral goal. A total of 30 control subjects underwent the same health and behavioral assessments at intake and program end but did not receive coaching or SMS text messages.

Results: The texting app showed that 87.31% (2187/2505) of messages were viewed by intervention participants. Furthermore, 28 of the 30 intervention participants and all 30 control participants provided outcome data. Among intervention participants, 22 of 30 (73%) showed improvement in health behavior goal attainment, with the whole group (n=30) showing a mean improvement of 88% (95% CI 39-136). Mean improvement in any behavioral domains was not seen in the control group. Intervention participants also increased their exercise significantly compared with controls, regardless of their self-selected goal category. The increased exercise was paralleled by significantly lower fasting glucose levels.

Conclusions: The health coaching session plus tailored SMS text messages improved self-selected health behaviors with a modest ripple effect to include unselected health behaviors.

Trial Registration: Clinicaltrials.gov NCT02476604; https://clinicaltrials.gov/ct2/show/NCT02476604 (Archived by WebCite at http://www.webcitation.org/6qAAryS5t)

(JMIR Mhealth Uhealth 2017;5(5):e67) doi:10.2196/mhealth.6638

KEYWORDS

health behaviors; diet habits; exercise; sleep; text telecommunications; universities

Introduction

During the transition to campus life, college students face challenges related to eating behaviors, physical activity, stress, and sleep [1-4]. Numerous factors may contribute to unhealthy lifestyles among students [5]. The all-you-can-eat dining facilities, time constraints, and lack of food preparation facilities in the dormitories affect food choices [5]. Diminished physical activity [6] and heightened levels of perceived stress [2], especially in women [7], can influence weight. It has been well documented that college students experience significant weight gain and changes in body composition, notably in their freshman year, putting them at higher risk for chronic diseases [8-10]. A need exists for effective strategies to encourage college students to adopt and maintain healthy lifestyles.

Mobile technology is commonplace today, particularly on college campuses [11]. As defined by a recent review, mobile technology broadly includes bidirectional message-type activities via conventional mobile phones, smartphones, and portable laptop computers that allow for exchange of text messages (short message service, SMS), electronic mail, live video feeds, and Internet access [11]. Because nearly all students have cell phones with them most of their waking time [11], text message–based interventions represent a promising opportunity to offer remote health coaching to students on the go [12,13]. Since the introduction of mobile phone–based interventions to promote behavior change, numerous researchers have reported positive outcomes [13]. For example, research has investigated the efficacy of mobile phone–based interventions on smoking cessation [14-17], weight loss [18-21], physical activity [22-25], and improving clinical outcomes of pregnancy [26-28]. A mobile phone–based intervention that encourages healthy changes in a variety of lifestyle behaviors would have clear advantages, offering efficiency as several health behaviors could be addressed through a single app. However, most mobile phone–based interventions have focused on a single health objective to date.

This preliminary randomized controlled trial aimed to evaluate the effect of a single coaching session bolstered by tailored text messages on self-selected health goals (described in further detail in the Methods section) in a sample of college students. We hypothesized that the coaching plus text message–based intervention would result in students successfully attaining their specified goals in the intervention group, which was randomly selected from interested participants. In this study, we also performed a comparison with a control group to determine whether or not the goal-oriented improvements would be associated with improvements in other lifestyle behaviors to benefit the students' health in areas beyond their self-selected area of focus. The study also incorporated objective tests (serum glucose and cholesterol) that may be responsive to improved lifestyle behaviors.

Methods

Participants

This study was conducted at a small university in western Pennsylvania from January 2015 to May 2015 in collaboration with the Integrative Cardiac Health Project, Walter Reed Military Medical Center in Bethesda, Maryland. The university Institutional Review Board and the Human Research Protection Office of the US Army Medical Research and Materiel Command approved the study protocol for human subject research. The study protocol was registered on ClinicalTrials.gov with trial number 375278-3 (ClinicalTrials.gov Identifier: NCT02476604). Written informed consent was obtained from eligible participants before enrolment. The full trial protocol may be obtained by written request to the corresponding author.

Participants were recruited using various strategies such as classroom announcements, posted advertisements, flyers, and booths located in student common areas on the university campus. Students were eligible to participate if they were full-time students at the university and did not have any exclusion criteria. The exclusion criteria included (1) age less than 18 years and greater than 30 years; (2) pregnancy, suspected pregnancy, or plans to become pregnant; (3) diagnosis of a chronic medical condition; (4) history of an eating disorder; and (5) taking any prescribed medication other than an oral contraceptive.

Data Collection

A comprehensive health evaluation was performed by certified health coaches at enrolment in the study. Health coach certification was conferred by the American Council on Exercise, an accredited body that teaches skills for active listening and providing motivational feedback. The coaching skill set is applicable for a variety of health topics including diet, exercise, stress management, and sleep habits. The comprehensive evaluation upon enrolment in the study included a focused physical assessment and collection of information such as past medical history, current medications, and a review of systems using typical demographic forms. At baseline and following the 8-week intervention period, participants completed validated behavioral surveys (described in detail below) assessing dietary habits, exercise activity, levels of perceived stress, and sleep practices. Blood testing for fasting glucose and cholesterol was also performed before and after the intervention (details below).

Behavioral Surveys

(1) Rate Your Plate [29] (RYP) dietary assessment: The RYP consists of 27 questions with responses from 1 to 3 yielding a scoring range from 27 to 81. A higher RYP score indicates a more healthful eating pattern. Convergent validity for the RYP was validated in another setting in a comparison study with a previously validated survey tool, the Willett food frequency questionnaire [30]. Using the Pearson product-moment correlation, the two questionnaires showed a strong degree of correlation with r=.4-.5 (P<.001), on a range of specific dietary variables [29].

(2) International Physical Activity Questionnaire [31] (IPAQ): The IPAQ asks respondents to recall their physical activity levels for a "usual week" and relates the recalled activities to metabolic equivalents (METs) for standardization of data. The IPAQ has been validated across 12 nations as an exercise survey questionnaire including college-aged adults to be self-administered or administered by interview. The IPAQ provides reproducible data (test-retest reliability, concurrent validity, and criterion IPAQ validity against accelerometry) with a Spearman rho of .8 [31].

(3) Perceived Stress Scale [32] (PSS 14): This 14-item questionnaire asks the respondent how often certain experiences of stress occurred in the last month and is designed to measure the degree to which situations in one's life are appraised as stressful. With item responses from 0 to 4, the range of possible scores is 0 to 56, with higher scores correlating with higher stress. The PSS is designed for use in community samples with at least a junior high school education. The items are easy to understand and the response alternatives are simple to grasp. Moreover, the questions are quite general in nature and hence relatively free of content specific to any subpopulation group. Scores in the low 20s reveal modest stress levels, whereas scores approaching 30 are substantial and concerning. A validation study [32] shows that the PSS 14 has an internal consistency reliability of .85 by Cronbach alpha and a test-retest reliability of .85.

(4) Pittsburgh Sleep Quality Index [33] (PSQI): The PSQI is a self-rated questionnaire that assesses sleep quality and disturbances over a 1-month time interval. Nineteen individual items generate 7 component scores whose sum yields a global score with a range of 0 to 21. A global score of greater than 5 indicates a poor sleeper. Sleep perturbations can be categorized by scores: 0 to 5 is a good sleep score; 6 to 10 shows mild sleep difficulty; 11 to 15 moderate sleep difficulty; and 16 to 21 severe sleep difficulty. The psychometric and clinical properties of the PSQI suggest its utility both in clinical practice and research activities. A PSQI greater than 5 has a diagnostic sensitivity of 90% and specificity of 87% (kappa > .75, P<.001) [33].

(5) Epworth Sleepiness Scale [34] (ESS): The ESS is the most widely used tool to estimate the subjective symptom of daytime sleepiness. Respondents are asked to use a scale of 0 to 3 to estimate their likelihood of dozing in 8 different situations in recent weeks. The individual scores are summed and possible scores range from 0 to 24. A score greater than 10 indicates excessive daytime sleepiness. The ESS has a sensitivity of 94% and a specificity of 100% for detecting excessive daytime sleepiness [34].

(6) Visual Analog Fatigue Scale [35] (FS): The FS was developed by the Stanford Patient Education Research Center. The scale was tested with 122 individuals, with a mean value of 4.9 and standard deviation of 2.7. This fatigue scale asks respondents to express their experience of fatigue from 0 to 10 for the previous 2-week period. Individuals denoting scores of 0 to 4 have no-to-minimal fatigue, those who circle 5 to 7 express milder fatigue, and those selecting 8 to 10 experience more severe fatigue.

Blood Collection

A certified phlebotomist collected two blood samples from each participant, one sample upon enrolment in the study and the second sample at the end of the 8-week study period. Participants in both the intervention arm and the control arm received a US $50.00 payment for each of the two blood samples taken. Participants were asked to fast for 12 hours before blood collection. All blood collections occurred between 7:00 AM and 9:00 AM. Serum samples were collected in vacutainers and separated by centrifuging at 4150 rpm for 5 minutes using a Sorvall Legend Mach 1.6 centrifuge (Thermo Fisher Scientific Inc). Samples were stored at 22°C and analyzed for total cholesterol and glucose at a regional commercial laboratory.

Intervention

The participants were randomly assigned in a one-to-one ratio by a computer program to one of two groups for the duration of the study. As part of the informed consent procedure, participants were made aware that the study involved a flip-of-the-coin chance that they may experience different components of the study compared with another participant. After randomization, participants were further informed regarding the particulars of their expected experience. One group received the intervention of behavioral goal setting with a health coach (certified by the American Council on Exercise) followed by text messaging throughout the 8-week study period (described in detail below). The other group served as controls, receiving their baseline survey results and blood test results in writing but not setting behavioral goals with the health coach and not receiving text messages during the study period.

Table 1. Examples of behavioral goals with intervention messages.

Goals and messages	Nutrition	Exercise	Stress	Sleep
Goal	Increase fruit and vegetable intake from 3 per day to 5 per day.	Increase exercise from 0 to 3 times per week for 20 minutes. (60 minutes per week)	Make a "to-do" list 5 days per week instead of 0 days per week.	Make bedtime 11:00 PM instead of midnight.
Example messages				
Action-oriented	Try to add two servings of fruit to your day, a handful of grapes, a crisp apple, etc. Natural fruit is a great source of fiber, provides fullness and sweetness.	Try this workout: 5 minutes of a light jog, 1 minute of sprinting followed by 1 minute of low speed jogging, repeat 5 times. Jog for 5 minutes as your cool down.	Try to organize for 15 minutes each day. This could mean anything from sorting mail to throwing out mystery foods in the refrigerator. Just 15 minutes per day can make a huge difference over time! Enjoy your weekend!	Did you know that stretching before bed could help you sleep better? Give this stretch a try: Sit on floor with pillow in front of you. Bend left knee, bringing bottom of left foot to right inner thigh. Lift butt and extend right leg behind you. Stay centered and lean forward from hips, placing head on pillow. Extend arms forward, elbows are slightly bent. Hold for 8-10 breaths. Roll back up. Switch legs; repeat.
Motivational	Good job with your fruits and veggies! Looks like you've improved!	No matter how slow you go. You are still lapping everyone on the couch.	Make each day count. We don't have all the time in the world. So focus on today and do the things you really want to do. :)	Perform at your highest potential. Sleep prepares your body and mind for the next day's endeavors.
Informational	Too much salt in our diet can raise blood pressure. Most fruits and vegetables are naturally low in salt and can help improve blood pressure. Focus on putting more fresh food like fruits and vegetables into your meals.	Exercise releases endorphins, which creates feelings of happiness and euphoria.	I heard the app "Simplenotes" is a good one for your iPhone for organization and it's free...you should look into it!	Did you know? Watching television before bedtime can actually stimulate the mind rather than relax it. Listening to audiobooks or music may be a better choice.

Following baseline data collection, results were aggregated so that study participants could compare their results to normative standards. This allowed participants to visualize potential areas of their health behaviors that would benefit from improvement. Participants were sent their health assessment to review in preparation for a meeting in person with the health coach. During the face-to-face meeting with the health coach, participants in the intervention group were asked to set one behavioral goal in a domain of diet, exercise, stress management, or sleep. The behavioral goal represented an actionable behavior that the participant could work to improve. (See Table 1 for sample goals.) Weight loss was not included as a goal because weight loss itself is a result of behavioral change, not a behavior itself. Health coaches also used the face-to-face interaction to learn about the participant's routine, habits, and motivational cues providing relevant information for tailoring messages. This initial face-to-face meeting required 45 minutes to 1 hour to complete.

SMS Text Messages

Before the creation of text message content, focus groups were conducted in order to understand preferences of the college student target population (A Roth, unpublished data, September, 2014). The dialogue captured during these focus groups suggested that college students preferred a mix of action-oriented, informational, and motivational messages that were succinct and sent no more than three times per week (Table 1). The investigators elected not to exceed the frequency of 3 messages per week so as to avoid turning an encouraging communication into an aggravation that would be quickly deleted. The health coaches disseminated intervention SMS text messages through a custom iOS application designed by programmers working with the Integrative Cardiac Health Project. Study participants downloaded the iOS application to a university-issued iPad or a personal iPhone. Intervention messages were delivered on a regular schedule on Tuesdays and Thursdays at 9:00 AM and Saturdays at 11:00 AM for the study period of 8 weeks. The schedule for delivery of messages was chosen to distribute the messages across the week at times that would not be disruptive of the students' living schedule. Students retained the freedom to check for messages at their convenience. There were two health coaches hired to conduct the study. Both of the coaches worked in equal measure with participants who were randomized to the intervention arm and with those who were randomized to the control arm. Once a participant was assigned to one of the health coaches, that participant received all contact from that particular health coach for the duration of the study. Health coaches were in contact with the participants randomized to the control arm at the beginning of the study and again at study completion. The health coaches did not have contact with the control participants during the 8 weeks of the study period. Both coaches were certified

for training in the same health coaching techniques and provided text messages to the participants randomized to the intervention arm using guidance from the participants themselves as to what sort of messages would help them stay on track to reach their self-selected health behavior goal. Participants in the intervention arm were asked weekly via the application to complete brief behavioral assessments (BA). The BA included a question addressing progress toward the self-selected goal. The coaches used BA feedback to customize text messages to address barriers impeding goal attainment. Some participants asked for more motivational type of text messages, whereas other participants opted for action-oriented messages that instructed them on their behaviors. A third type of message preferred by some participants was an informational-type text message. (See Table 1 for sample messages.) The coaches required approximately one hour for each intervention participant per week to read the brief BA feedback messages and to customize the text messages to be disseminated to each participant. The self-selected behavioral goal served as a basis for text messages designed to encourage goal attainment. Half of the messages sent as part of the intervention were directed toward the self-selected behavioral goal and half of the messages targeted other health behaviors more generally. Supervision of the text message development was provided by the study's principal investigator.

Sample Size Estimation

The RYP dietary assessment tool served as an outcome measure to determine the sample size estimation because the investigators had a past experience with the RYP tool and could estimate an effect size of the intervention on this parameter. It was assumed that similar responsiveness in the other outcome measures would be seen and that by using a general "improvement score" for each behavioral dimension, participants could be offered the freedom to choose any of the four behaviors for concentration of their improvement efforts. Dietary changes in RYP were expected to average 10 (SD 2) points for the intervention group compared with zero change (same standard deviation) for the control group. With a power of 90% for rejecting the null hypothesis and using a one-sided test with alpha = 2.5% (as we expected an increase in dietary adherence and minimal chance of a decrease in adherence), the sample size was found to be 25 participants in each arm of the study. Accounting for a 30% dropout rate, a total of 70 participants would be recruited, 35 participants for each arm.

Statistical Methods

Statistical analysis was performed using SPSS for Macintosh version 22.0.0.0 (IBM, 2013). Descriptive statistics were calculated for variables including means, standard deviations, medians, and minimum and maximum values. Skewness and kurtosis values of the dependent variables were analyzed for significant deviation from a normal distribution. Further comparisons between participants randomized to the intervention group and those serving as controls were made using two-sample *t* tests or chi-square tests as appropriate, with a *P* value <.05 indicating statistical significance.

Because participants set unique behavioral goals, we used a standardized indicator that operationalized "success" as the percentage of each participant's goal that was attained at the conclusion of the study. For example, at intake one participant completed 7758 steps per day. She set a goal of 10,000 steps per day, and by the completion of the study, she had completed 9008 steps per day. Thus, she attained 56% (ie, 100 × [9008−7758]/[10,000−7758]) of her goal. The intervention would be considered successful if the mean percentage of goal attainment was statistically significantly greater than 0.

Results

General

A total of 84 potential participants were screened for eligibility to participate in the study. Following baseline data collection, 3 participants were released due to prescription medication usage and 1 because baseline blood pressure results showed hypertension deserving medical attention. Other attrition factors included withdrawing from the university, change in student status, electing not to participate in the study due to time constraints, or other personal reasons (Figure 1). Of 60 enrolled participants, 30 were randomized to the intervention arm and 30 to the control arm. For 2 participants in the intervention arm, end-of-study data were not obtained to determine success or failure of the intervention with respect to goal attainment. In order to maintain an intention-to-treat analysis, these 2 participants were assigned zero progress toward their behavioral goals and their data were retained in the final analysis.

Demographic information on the 60 participants is presented in Table 2. The study population comprised young adults, predominantly white (87%), and the majority were women (68%). A majority of participants were freshmen (43%) or sophomores (30%), and most participants (80%) lived on campus during the study period. As shown in Table 2, there were no differences in demographic characteristics between participants randomized to the intervention and control arms.

Table 2. Demographic Information.

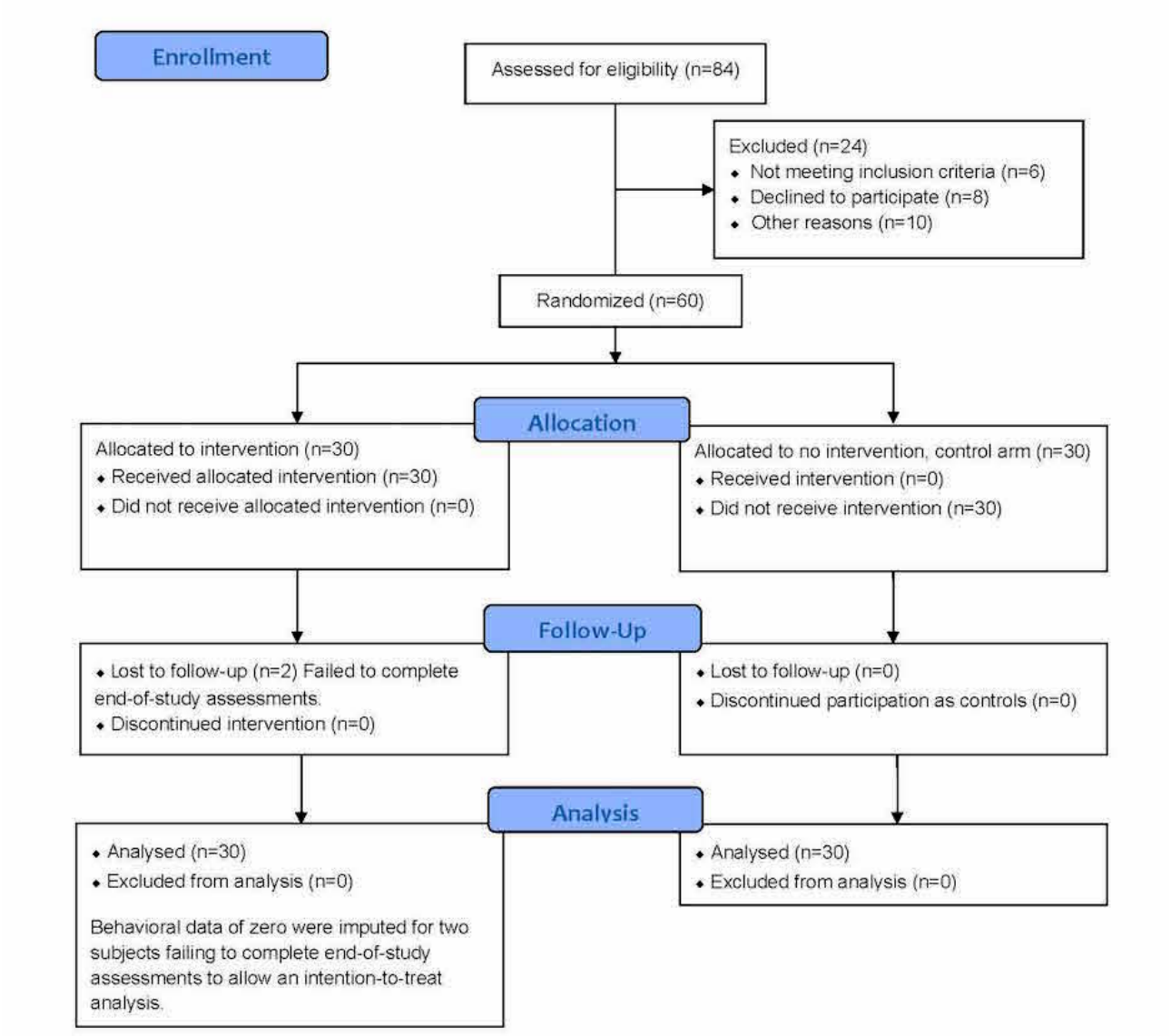
Demographic characteristics	All participants (n=60)	Control (n=30)	Intervention (n=30)	*P* value[a]
Age in years, mean (SD)	19.4 (1.0)	19.3 (0.9)	19.5 (1.1)	.38
Gender (women), n (%)	41(68)	20 (67)	21 (70)	.79
Race[b]	52W, 5B, 2H, 1A	24W, 4B, 1H, 1A	28W, 1B, 1H	.38
Grade point average, mean (SD)	3.5 (0.4)	3.5 (0.4)	3.6 (0.3)	.16
Year of study[c]	26F, 18So, 9J, 7Sr	12F, 11So, 5J, 2Sr	14F, 7So, 4J, 5Sr	.49
On/off campus residence	48/12	24/6	24/6	>.99
Active in varsity sport, n (%)	27 (45)	16 (53)	11 (37)	.42

[a]*P* value signifies level of statistically significant differences by *t* test between control and intervention subjects.

[b]W=white, B=black, H=Hispanic, A=Asian.

[c]F=freshman, So = sophomore, J=junior, Sr = senior.

Figure 1. CONSORT flow diagram.



Goal Attainment for Intervention Participants

The computer server was queried to compile information on messaging. During the course of the study, 1160 messages were sent to intervention participants and 982 messages (85%) were viewed.

Values for goal attainment were acceptable for a normal distribution (D_{27}=.95, *P*=.19; skewness=0.28, SE=0.45; and kurtosis=1.18, SE=0.87). Of the 30 participants who received the intervention, 22 (73%) showed improvements in their self-declared behavioral goal, 6 of 30 (20%) did not make gains toward their behavioral goal, and 2 (7%) did not provide data that could be evaluated. For analysis, the 2 participants who failed to complete end-of-study assessments were assigned zero values for behavioral goal attainment and were therefore considered failures of the intervention. For the entire intervention group (n=30), the success rate of 88% was calculated by averaging the percentages of each participant's goal (positive gains, negative, and zeros) that was attained at the conclusion of the study. For the 6 participants who did not make gains toward their goals, 1 reported no change and 5 reported negative movement, that is, movement away from their goals, averaging negative 66%. A one-sample *t* test on participants' percent of goal attainment scores was used to determine whether or not their mean was significantly greater than 0 (a score of 0 or less would indicate that the intervention was not effective because the participant would have achieved 0% of his or her goal or, indeed, had regressed on the goal indicator following the intervention). The one-sample *t* test indicated that the percent of goal attainment of 88% (for n=30) was significantly greater than 0 (t_{26}=3.68, *P*<.001, Cohen d=0.71, 95% CI [38.56-136.11]). Our hypothesis that the coaching plus text intervention would help participants' goal attainment was supported by the data.

Additionally, data were examined to determine whether or not the participants' success depended upon the category of their self-selected behavioral goal. Of the 30 participants in the intervention group, 10 selected dietary goals, 8 selected exercise goals, whereas there were 6 participants each who selected goals in stress management and in sleep. Based upon the category of behavioral goal, there were no differences in means for percentage of goal attainment ($F_{3,20}$=.16, *P*=.92, partial eta-squared = .29). The breakdown in the percentages of goal attainment across the 4 categories was diet (131%), exercise (75%), stress management (69%), and sleep (98%).

Since the control participants did not establish behavioral goals, it was not possible to measure control group progress toward a composite behavioral goal at the outcome of the study as was done for the intervention participants. However, the validated behavioral surveys were completed by control participants at baseline and at the outcome date of the study. As was anticipated, control participants did not show statistically significant changes in any of the behavioral domains (for diet *P*=.55, for exercise *P*=.77, for stress management *P*=.75, and for sleep *P*=.93).

Broader Effects of a Self-Selected Behavior Goal

Upon comparing intervention participants with the control group, irrespective of their particular behavioral goals, it was found that exercise was the only behavior difference, with intervention participants increasing their reported exercise more than controls to a statistically significant degree (Table 3). Intervention participants increased their exercise by 19% compared with control participants, who decreased their exercise by 6%. There were no other differences in health behaviors or symptoms between intervention participants and controls, implying that there was only a modest "ripple effect" to produce behavioral improvements outside of the self-selected behavioral goal.

Participants receiving the intervention showed a statistically significant improvement in fasting blood glucose compared with controls (6% decrease vs 1% decrease), while cholesterol levels were not different (Table 3).

Table 3. Lifestyle behaviors, symptoms, and laboratory data at program baseline and program completion.

Behavior or Objective Measure	Program baseline				Program completion			
	All participants (n=60)	Control (n=30)	Intervention (n=30)	*P* value[a]	All participants (n=60)	Control (n=30)	Intervention (n=30)	*P* value[a]
Diet (RYP)[b]	61.4 (SD 7.8)	61.5 (SD 7.7)	61.3 (SD 8.1)	.91	63.0 (SD 8.9)	62.7 (SD 8.3)	63.6 (SD 9.6)	.81
Exercise (IPAQ)[c]	2416 (SD 2508)	2208 (SD 2106)	2632 (SD 2889)	.52	2590 (SD 1965)	2074 (SD 1410)	3144 (SD 2324)	.04
Stress (PSS-14)[d]	23.1 (SD 8.0)	23.1 (SD 8.9)	23.0 (SD 7.0)	.99	23.9 (SD 8.7)	23.8 (SD 9.3)	24.0 (SD 8.1)	.94
Sleep (PSQI)[e]	5.2 (SD 2.7)	5.3 (SD 2.3)	5.1 (SD 3.0)	.81	5.5 (SD 3.0)	5.9 (SD 3.2)	5.2 (SD 2.9)	.37
Sleepiness (ESS)[f]	7.8 (SD 3.8)	8.1 (SD 4.3)	7.5 (SD 3.2)	.55	8.5 (SD 4.0)	8.8 (SD 3.8)	8.3 (SD 4.2)	.63
Fatigue (SFS)[g]	4.1 (SD 2.0)	3.7 (SD 2.2)	4.4 (SD 1.9)	.21	4.6 (SD 2.0)	4.3 (SD 2.0)	4.9 (SD 2.0)	.25
Fasting glucose[h]	78.8 (SD 6.4)	79.0 (SD 6.7)	78.6 (SD 6.3)	.83	76.1 (SD 8.4)	78.2 (SD 8.2)	74.0 (SD 8.3)	.05
Fasting cholesterol[h]	159.4 (SD 30.1)	157.1 (SD 28.5)	161.7 (SD 32.0)	.56	159.7 (SD 29.1)	160.5 (SD 29.3)	158.9 (SD 29.4)	.83

[a]*P* value signifies level of statistically significant differences by *t* test between control and intervention subjects.

[b]RYP: Rate Your Plate Dietary Questionnaire; of the 81 total possible points, higher score indicates a healthier diet.

[c]IPAQ: International Physical Activity Questionnaire in metabolic equivalent-minutes per week, higher score equates with greater activity level.

[d]PSS-14: Perceived Stress Scale, of 56 total possible points, higher score indicates greater level of perceived stress.

[e]PSQI: Pittsburgh Sleep Quality Index, of 21 possible points, lower score indicates better sleep quality.

[f]ESS: Epworth Sleepiness Scale, of 24 possible points, higher score indicates greater daytime sleepiness.

[g]SFS: Stanford Fatigue Scale, of 10 possible points, higher score indicates greater severity of fatigue symptom.

[h]Fasting glucose and fasting cholesterol in mg/dL.

Discussion

Principal Findings

The main research hypothesis of this study was that a single counseling session followed by a text message–based intervention would result in the attainment of a self-selected health behavior goal. A secondary analysis was performed to determine whether or not the intervention directed at a particular behavior might also have a broader effect by encouraging other healthful behavior changes outside of the behavior goal of choice.

Similar to many previous studies [12-28], this study's findings showed that the coaching session bolstered by tailored text messages led to significant within-group behavior change for the self-selected goal. These behavior changes were substantial in magnitude. The findings also suggest that there was a small ripple effect to stimulate improvement in exercise health behaviors even when exercise was not the self-selected behavior goal.

The importance of the improved exercise habits is underscored by the clinically relevant improvement in fasting glucose levels in the intervention group compared with the control group. These objective data were validating for the subjective self-reported exercise levels. Although the absolute decrease in serum glucose values appears to be modest (on average 4.6 mg/dL for the intervention participants), there is epidemiological evidence supporting the clinical relevance of this amount of glucose lowering when sustained over time. A recent prospective study in healthy individuals who did not have diabetes or cardiovascular disease showed that each 3 mg/dL increase in fasting glucose, even in the clinically accepted normal range, predicted a 1% increased risk of developing coronary disease over a mean of 4.3 years [36]. This finding supports the observation that even small improvements in health behaviors begun at a young age can produce major health improvements in the long term.

Researchers conducting numerous systematic reviews and meta-analyses analyzing text message–based studies have concluded that despite findings that text message interventions do work to encourage adherence to therapy, there are

methodological issues with research in this area [13,15,18,25,37,38]. The current protocol aimed to address several of these issues. For example, the attributes of text messages such as content and dosage are often not described or evaluated to determine how those factors influence the results [18,25,37]. Additionally, text messages are sometimes standardized and generic for mass audiences or targeted for a specific health topic [38] (ie, smoking cessation, weight loss). It has been suggested that the tailoring of messages to a specific individual can enhance effect [17,39,40]. The current protocol provided a customized approach in which (1) participants could self-select an area of their health they elected to address and (2) the messages they received were tailored in response to preparatory work with focus groups and by feedback they sent to their coaches.

Another methodological design issue that the current protocol was designed to address was the use of objective data collection to evaluate outcome measures in tandem with subjective self-report progress toward their goal. To this end, fasting glucose values in the intervention group were significantly lower compared with the control group, in parallel with the increased exercise levels attained.

Limitations

There are limitations to this study. One limitation in the study design is that the intervention comprised two components: (1) a single counseling session with goal setting and (2) an 8-week period of SMS text message support. It is therefore not possible to determine which of these components singly or together was responsible for the positive effects of the intervention. Limiting the intervention to one of these components would have enabled the investigators to report the effect of counseling or of text messaging. Future research on text messaging would benefit from unbundling such an intervention, perhaps by providing the counseling session to both the intervention and the control arms of the study. By setting the intervention to a single component such as text message support, it would be feasible to investigate the different effects of generic text messages versus texts with tailored, personalized feedback. Such an investigation is worthy of attention because generic texts can be automated for distribution, a much less time-consuming process than providing texts with personalized content.

The study did have statistical power to demonstrate a significant change in overall healthy behaviors in response to the intervention. However, the study may have been underpowered to find significant changes for the different individually selected behaviors with sample sizes of 6-10 participants. The small sample sizes likely explain why none of the analyses for separate behaviors were significant.

Another limitation of this report is the lack of analysis of the reports received by the health coaches for weekly behavioral compliance of the participants. This lack of analysis stemmed from the decision to specify *a priori* an overall 8-week outcome to the intervention. However, future studies may find this information helpful to interpret how well the individualized or customized the text messages worked. It may also be useful to interpret the overall burden of the intervention as experienced by the participants.

The study intervention lasted only 8 weeks, a trial period that was insufficient to capture significant improvements in some measures such as cholesterol changes. With this short study duration, it was also not possible to determine whether or not behavioral gains would be maintained over longer periods.

By design and in order to avoid confounding variables, the study excluded participation by students who had chronic illnesses or who took prescription medications other than for birth control. Furthermore, the enrolled population was predominantly white women. These factors do limit the ability to generalize the findings of this study to the student population at large.

Conclusions

The findings from this SHUPEP study do support the current body of literature evaluating the utility and effectiveness of tailored mobile-based health communications as a successful strategy for improving health behaviors [17-39]. A new finding from this study suggests that there may be a mild ripple effect of this technological intervention on other health outcomes outside of the targeted behavior. Perhaps the main takeaway message for care providers on other college campuses is that the face-to-face feedback and behavioral health goal-setting along with timely motivational text messages as described in this report are of value in promoting healthy behaviors in college students. Care providers for college students should be emboldened to embrace digital technologies to augment their efforts for health promotion.

Acknowledgments

This research was funded by the US Army Medical Research and Materiel Command by a grant to the Henry M Jackson Foundation for the Advancement of Military Medicine, Rockville, Maryland, with Jackson Foundation Award Number: W81XWH-11-2-0178. The US Army Medical Research and Materiel Command and the Jackson Foundation for the Advancement of Military Medicine had no part in study design, data analysis, review, or approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT eHealth Checklist.

[PDF File (Adobe PDF File), 77KB - mhealth_v5i5e67_app1.pdf]

References

1. Nelson MC, Kocos R, Lytle LA, Perry CL. Understanding the perceived determinants of weight-related behaviors in late adolescence: a qualitative analysis among college youth. J Nutr Educ Behav 2009;41(4):287-292. [doi: 10.1016/j.jneb.2008.05.005] [Medline: 19508935]
2. Vella-Zarb RA, Elgar FJ. The 'freshman 5': a meta-analysis of weight gain in the freshman year of college. J Am Coll Health 2009;58(2):161-166. [doi: 10.1080/07448480903221392] [Medline: 19892653]
3. Ye L, Hutton JS, Keane K, Manasia M, Gregas M. Napping in college students and its relationship with nighttime sleep. J Am Coll Health 2015 Jan;63(2):88-97. [doi: 10.1080/07448481.2014.983926] [Medline: 25397662]
4. Eliasson A, Lettieri C, Eliasson A. Early to bed, early to rise! sleep habits and academic performance in college students. Sleep Breath Medline 2010;14(1):71-75. [Medline: 19603214]
5. Lacaille LJ, Dauner KN, Krambeer RJ, Pedersen J. Psychosocial and environmental determinants of eating behaviors, physical activity, and weight change among college students: a qualitative analysis. J Am Coll Health 2011;59(6):531-538. [doi: 10.1080/07448481.2010.523855] [Medline: 21660808]
6. Butler SM, Black DR, Blue CL, Gretebeck RJ. Change in diet, physical activity, and body weight in female college freshman. Am J Health Behav 2004;28(1):24-32. [Medline: 14977156]
7. Serlachius A, Hamer M, Wardle J. Stress and weight change in university students in the United Kingdom. Physiol Behav 2007 Nov 23;92(4):548-553. [doi: 10.1016/j.physbeh.2007.04.032] [Medline: 17537466]
8. Lloyd-Richardson EE, Bailey S, Fava JL, Wing R, Tobacco Etiology Research Network (TERN). A prospective study of weight gain during the college freshman and sophomore years. Prev Med 2009 Mar;48(3):256-261 [FREE Full text] [doi: 10.1016/j.ypmed.2008.12.009] [Medline: 19146870]
9. Gropper S, Simmons K, Connell L, Ulrich P. Changes in body weight, composition, and shape: a 4-year study of college students. Appl Physiol Nutr Metab 2012 Dec;37(6):1118-1123. [doi: 10.1139/h2012-139] [Medline: 22978391]
10. Holm-Denoma JM, Joiner TE, Vohs KD, Heatherton TF. The “freshman fifteen” (the “freshman five” actually): predictors and possible explanations. Health Psychol 2008 Jan;27(1 Suppl):S3-S9. [doi: 10.1037/0278-6133.27.1.S3] [Medline: 18248103]
11. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
12. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
13. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
14. Haug S, Schaub M, Venzin V, Meyer C, John U. Efficacy of a text message-based smoking cessation intervention for young people: a cluster randomized controlled trial. J Med Internet Res 2013 Aug 16;15(8):e171 [FREE Full text] [doi: 10.2196/jmir.2636] [Medline: 23956024]
15. Spohr SA, Nandy R, Gandhiraj D, Vemulapalli A, Anne S, Walters ST. Efficacy of SMS text message interventions for smoking cessation: a meta-analysis. J Subst Abuse Treat 2015 Sep;56:1-10. [doi: 10.1016/j.jsat.2015.01.011] [Medline: 25720333]
16. Bock B, Heron K, Jennings E, Morrow K, Cobb V, Magee J, et al. A text message delivered smoking cessation intervention: the initial trial of TXT-2-quit: randomized controlled trial. JMIR Mhealth Uhealth 2013 Jul;1(2):e17 [FREE Full text] [doi: 10.2196/mhealth.2522] [Medline: 25098502]
17. Skov-Ettrup LS, Ringgaard LW, Dalum P, Flensborg-Madsen T, Thygesen LC, Tolstrup JS. Comparing tailored and untailored text messages for smoking cessation: a randomized controlled trial among adolescent and young adult smokers. Health Educ Res 2014 Apr;29(2):195-205 [FREE Full text] [doi: 10.1093/her/cyt112] [Medline: 24399268]
18. Siopis G, Chey T, Allman-Farinelli M. A systematic review and meta-analysis of interventions for weight management using text messaging. J Hum Nutr Diet 2015 Feb;28(Suppl 2):1-15. [doi: 10.1111/jhn.12207] [Medline: 24480032]
19. Haapala I, Barengo NC, Biggs S, Surakka L, Manninen P. Weight loss by mobile phone: a 1-year effectiveness study. Public Health Nutr 2009 Dec;12(12):2382-2391. [doi: 10.1017/S1368980009005230] [Medline: 19323865]
20. Steinberg DM, Levine EL, Askew S, Foley P, Bennett GG. Daily text messaging for weight control among racial and ethnic minority women: randomized controlled pilot study. J Med Internet Res 2013;15(11):e244 [FREE Full text] [doi: 10.2196/jmir.2844] [Medline: 24246427]
21. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009 Jan;11(1):e1 [FREE Full text] [doi: 10.2196/jmir.1100] [Medline: 19141433]
22. Prestwich A, Perugini M, Hurling R. Can implementation intentions and text messages promote brisk walking? a randomized trial. Health Psychol 2010 Jan;29(1):40-49. [doi: 10.1037/a0016993] [Medline: 20063934]

23. Maddison R, Pfaeffli L, Stewart R, Kerr A, Jiang Y, Rawstorn J, et al. The HEART mobile phone trial: the partial mediating effects of self-efficacy on physical activity among cardiac patients. Front Public Health 2014;2:56 [FREE Full text] [doi: 10.3389/fpubh.2014.00056] [Medline: 24904918]
24. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007 Apr;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
25. Buchholz S, Wilbur J, Ingram D, Fogg L. Physical activity text messaging interventions in adults: a systematic review. Worldviews Evid Based Nurs 2013 Aug;10(3):163-173. [doi: 10.1111/wvn.12002] [Medline: 23746267]
26. Evans WD, Wallace JL, Snider J. Pilot evaluation of the text4baby mobile health program. BMC Public Health 2012 Nov;12:1031 [FREE Full text] [doi: 10.1186/1471-2458-12-1031] [Medline: 23181985]
27. Evans WD, Wallace BJ, Szekely D, Nielsen P, Murray E, Abroms L, et al. Initial outcomes from a 4-week follow-up study of the Text4baby program in the military women's population: randomized controlled trial. J Med Internet Res 2014 May;16(5):e131 [FREE Full text] [doi: 10.2196/jmir.3297] [Medline: 24846909]
28. Soltani H, Duxbury AM, Arden MA, Dearden A, Furness PJ, Garland C. Maternal obesity management using mobile technology: a feasibility study to evaluate a text messaging based complex intervention during pregnancy. J Obes 2015;2015:814830 [FREE Full text] [doi: 10.1155/2015/814830] [Medline: 25960889]
29. Gans K, Sundaram S, McPhillips J, Hixson M, Linnan L, Carleton R. Rate your plate: An eating pattern assessment and educational tool used at cholesterol screening and education programs. J Nutr Educ Behav 1993 Jan;25(1):29-36. [doi: 10.1016/S0022-3182(12)80186-5]
30. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985 Jul;122(1):51-65. [Medline: 4014201]
31. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutr 2006 Sep;9(6):755-762. [Medline: 16925881]
32. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
33. Buysse D, Reynolds C, Monk T, Berman S, Kupfer D. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [Medline: 2748771]
34. Johns MW. Reliability and factor analysis of the Epworth sleepiness scale. Sleep 1992 Aug;15(4):376-381. [Medline: 1519015]
35. Stanford Patient Education Research Center. 2016. Fatigue visual numeric scale URL: http://patienteducation.stanford.edu/research/vnsfatigue.html [accessed 2016-09-03] [WebCite Cache ID 6kFjI93wn]
36. Shaye K, Amir T, Shlomo S, Yechezkel S. Fasting glucose levels within the high normal range predict cardiovascular outcome. Am Heart J 2012 Jul;164(1):111-116 [FREE Full text] [doi: 10.1016/j.ahj.2012.03.023] [Medline: 22795290]
37. U.S. Department of HealthHuman Services. Mathematica-mpr. 2014. Using health text messages to improve consumer health knowledge, behaviors, and outcomes: an environmental scan URL: https://www.mathematica-mpr.com/our-publications-and-findings/publications/using-health-text-messges-to-improve-consumer-health-knowledge-behaviors-and-outcomes [accessed 2017-05-11] [WebCite Cache ID 6qNKHMtcB]
38. 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]
39. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
40. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. Health Educ Res 2005 Oct;20(5):527-539. [doi: 10.1093/her/cyh014] [Medline: 15701665]

Abbreviations

BA: behavioral assessments
ESS: Epworth Sleepiness Scale
FS: Visual Analog Fatigue Scale
IPAQ: International Physical Activity Questionnaire
MET: metabolic equivalent
PSQI: Pittsburgh Sleep Quality Index
PSS: Perceived Stress Scale
RYP: Rate Your Plate

Edited by G Eysenbach; submitted 10.09.16; peer-reviewed by K Heron, B Bock; comments to author 31.12.16; revised version received 04.03.17; accepted 09.03.17; published 17.05.17.

Please cite as:
Sandrick J, Tracy D, Eliasson A, Roth A, Bartel J, Simko M, Bowman T, Harouse-Bell K, Kashani M, Vernalis M
Effect of a Counseling Session Bolstered by Text Messaging on Self-Selected Health Behaviors in College Students: A Preliminary Randomized Controlled Trial
JMIR Mhealth Uhealth 2017;5(5):e67
URL: http://mhealth.jmir.org/2017/5/e67/
doi:10.2196/mhealth.6638
PMID:28526664

Original Paper

What to Build for Middle-Agers to Come? Attractive and Necessary Functions of Exercise-Promotion Mobile Phone Apps: A Cross-Sectional Study

Gen-Yih Liao[1,2], PhD; Yu-Tai Chien[3], MSc; Yu-Jen Chen[1], PhD; Hsiao-Fang Hsiung[4], PhD; Hsiao-Jung Chen[4], MSc; Meng-Hua Hsieh[1], MSc; Wen-Jie Wu[1,5], PhD

[1]Department of Information Management, Chang Gung University, Taoyuan City, Taiwan

[2]Dept of Nursing, Chang Gung Memorial Hospital, Taoyuan Branch, Taoyuan City, Taiwan

[3]Department of Information Management, National Taiwan University, Taipei, Taiwan

[4]Department of Nursing, Chang Gung University of Science of Technology, Guishan District, Taoyuan City, Taiwan

[5]Division of Chinese Gynecology, Center for Traditional Chinese Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan, Kwei-Shan Tao-Yuan, Taiwan

Corresponding Author:
Gen-Yih Liao, PhD
Department of Information Management
Chang Gung University
259 Wen-Hwa 1st Road
Guishan District
Taoyuan City, 333
Taiwan
Phone: 886 32118800 ext 5852
Fax: 886 32118020
Email: gyliao@acm.org

Abstract

Background: Physical activity is important for middle-agers to maintain health both in middle age and in old age. Although thousands of exercise-promotion mobile phone apps are available for download, current literature offers little understanding regarding which design features can enhance middle-aged adults' quality perception toward exercise-promotion apps and which factor may influence such perception.

Objectives: The aims of this study were to understand (1) which design features of exercise-promotion apps can enhance quality perception of middle-agers, (2) whether their needs are matched by current functions offered in app stores, and (3) whether physical activity (PA) and mobile phone self-efficacy (MPSE) influence quality perception.

Methods: A total of 105 middle-agers participated and filled out three scales: the International Physical Activity Questionnaire (IPAQ), the MPSE scale, and the need for design features questionnaire. The design features were developed based on the Coventry, Aberdeen, and London—Refined (CALO-RE) taxonomy. Following the Kano quality model, the need for design features questionnaire asked participants to classify design features into five categories: attractive, one-dimensional, must-be, indifferent, and reverse. The quality categorization was conducted based on a voting approach and the categorization results were compared with the findings of a prevalence study to realize whether needs match current availability. In total, 52 multinomial logistic regression models were analyzed to evaluate the effects of PA level and MPSE on quality perception of design features.

Results: The Kano analysis on the total sample revealed that visual demonstration of exercise instructions is the only attractive design feature, whereas the other 51 design features were perceived with indifference. Although examining quality perception by PA level, 21 features are recommended to low level, 6 features to medium level, but none to high-level PA. In contrast, high-level MPSE is recommended with 14 design features, medium level with 6 features, whereas low-level participants are recommended with 1 feature. The analysis suggests that the implementation of demanded features could be low, as the average prevalence of demanded design features is 20% (4.3/21). Surprisingly, social comparison and social support, most implemented features in current apps, were categorized into the *indifferent* category. The magnitude of effect is larger for MPSE because it effects quality perception of more design features than PA. Delving into the 52 regression models revealed that high MPSE more

likely induces *attractive* or *one- dimensional* categorization, suggesting the importance of technological self-efficacy on eHealth care promotion.

Conclusions: This study is the first to propose middle-agers' needs in relation to mobile phone exercise-promotion. In addition to the tailor-made recommendations, suggestions are offered to app designers to enhance the performance of persuasive features. An interesting finding on change of quality perception attributed to MPSE is proposed as future research.

(JMIR Mhealth Uhealth 2017;5(5):e65) doi:10.2196/mhealth.6233

KEYWORDS

physical exercise; middle aged; mobile application; self efficacy; consumer preference

Introduction

Background

Middle age begins when young adulthood ends and ends when old age starts [1]. The United Nation reported that the middle-age population in the more developed regions had increased dramatically from 2000 to 2010 and remain the greatest age group by 2024 [2], suggesting the importance of middle age studies. Because physiological functions (eg, lung function, muscle mass, renal blood flow, and bone density) begin to deteriorate in middle age, scientific guidelines suggest that physical exercise be taken to counter this deterioration effect [3]. Two studies examined the employees, aged 40-60 years, who changed their leisure time PA and found that increased PA reduced subsequent sickness absence and risk of disability retirement [4,5]. In addition, an empirical study has conducted a 14-year longitudinal research and indicated that middle-aged PA can reduce mortality and heart attacks in elder life [6]. Despite with the apparent importance of PA, middle-agers may not exercise as required. According to previous survey, 70% of middle-aged people did not meet exercise criteria defined by US federal government [7]. In addition, time spent sitting in middle-aged adults is suggested as too long [8]. As the characteristics of middle-aged life include established own family, clear career direction, and responsibility on children and aging parents [3], people in middle age are busy taking care of families and works, potentially decreasing their PA. Studies have also identified cost and time as barriers for middle-agers to adopt PA [3,9]. These findings suggest that it is important to intervene middle-aged sedentary lifestyle for elder health.

Intervention researchers are utilizing opportunities enabled by technology to design new health interventions. Tailoring an intervention and disseminating it using websites, by email or short message service (SMS) text messages, is considered as a promising health promotion strategy [10]. Recently, middle-agers are increasingly adopting mobile phones. In Nielsen's report [11] on the global mobile consumer released in 2013, the mobile phone penetration rates in middle-agers reached 40% or higher in some developed countries (eg, United States, United Kingdom, Italy, and South Korea). With the omnipresence and continuity of access, mobile phones therefore become an increasingly essential instrument for revolutionizing intervention strategies [12]. According to a meta-analysis [12], mobile phone apps in the category of PA promotion can measure sports statistics [13] and the number of steps [14], assist self-management (eg, functions with activity diary and reminders) [10]. However, variety of features unnecessarily guarantees the acceptance of users with information technology (IT)–enabled health care applications [15]. The unified theory of acceptance and use of technology (UTAUT) argues that users' expectations can influence their intention to accept mobile phone apps promoting PA in the context of a voluntary behavior [16]. The expectancy disconfirmation theory also suggests that, when prior-use expectations are matched or exceeded, users feel satisfied and their continuance intention toward product use will be increased [17]. Health care studies also suggest that meeting expectations is crucial not only for enhancing patient satisfaction [18], but also for relieving symptoms and reducing further use of health care resources [19]. Therefore, intervention designers and app developers would benefit in realizing what targeted users expect from using IT-based health care apps in order to deliver necessary functions [20].

To the best of our knowledge, Rabin and Bock [21] may be the first to examine user preferences regarding mobile phone apps related to PA. On the basis of 15 participants, their findings suggest that the most endorsed feature be automatic (and accurate) tracking of steps taken and calories burned, followed by visual tracking on exercise progress and several concrete functionalities (eg, body mass index [BMI] calculators) [21]. Due to the demographical characteristics of their study participants, however, the findings may not be applicable to other populations (eg, middle-agers). Furthermore, as the features were produced based on participant feedbacks, it is desirable to examine features based on a complete theoretical framework, which may contribute to comprehensive understandings both in the feature level and in the theory level.

Table 1. Empirical studies with the Kano method.

Authors (year)	Research domain	Product or service type	Research purpose	Product life stage
Chen and Chuang (2008) [26]	Technology	Mobile phone's body shape and button style	Product performance evaluation	Implementation or testing
Wang and Wu (2014) [27]	Technology	Mobile phone's core attributes (eg, CPU[a]) and optional attributes (eg, electronic wallet)	Feature classification	Prototype development
Palumbo, Dominici, and Basile (2013) [28]	Technology	Apps function (museum information [eg, opening or closing time], artworks [eg, photo]) and usability (friendly user interface)	Feature classification	Prototype development
Sulisworo and Maniquiz (2012) [29]	Health care	Registration, medical treatment, and physical facilities	Feature classification	Prototype development
Chang and Chang (2013) [30]	Health care	Physical facilities, staff characteristics, medical treatment, and administration	Service performance evaluation	Implementation or testing
Dominici and Palumbo (2013) [31]	Education	E-learning platform	Feature classification	Prototype development
Shahin and Zairi (2009) [32]	Airline	In-flight service, administration, and flight physical facilities	Feature classification	Prototype development

[a]CPU: central processing unit.

Aims of This Study

The aims of this study were to explore middle-aged adults' needs on functional features of mobile phone apps promoting PA. First, we created a representative set of 52 design features based on the Coventry, Aberdeen, and London—Refined (CALO-RE) taxonomy of 40 behavior change techniques (BCTs), as this taxonomy can improve the specification of PA interventions and strengthen the scientific study of intervention development [22] and has been applied to examine PA apps [23]. Second, to delve into middle-agers' needs, this study adopts the Kano method that interprets quality perception as a set of quality categories (eg, attractive, one-dimensional, indifferent, must-be, and reverse) that may influence customer satisfaction [24,25]. Table 1 presents previous studies adopting the Kano method to investigate user satisfaction toward product performance and evaluate the user perceptions of prototypes in design in various application domains (eg, mobile phone design, apps design, and health promotion), warranting applying the Kano method. With our study incorporated with the CALO-RE taxonomy and the Kano method, we can offer insights to realize users' perception on receiving comprehensive PA interventions via mobile phones and to predict their attitude toward using exercise-promoting apps.

To evaluate the intervention functions offered in mobile phone apps, previous research has applied BCTs to review 64 apps in iTunes and Google Play, and suggested that the apps included 5 BCTs on average [33]. A similar study also downloaded 100 top customer-rated PA apps in the "health and fitness" category of the Apple iTunes and Google Play and suggested that an average of 6.6 BCTs was used per app and most BCTs in the taxonomy were not represented in any app [34]. In addition, for each BCT, this study defines prevalence of a BCT as the percentage of apps implementing the BCT [34]. Although these studies offer insights into how prevalent BCTs are implemented in PA apps, user needs toward these persuasive features remain unexplored, indicating a research gap. Besides, as computer use self-efficacy and the habit of PA can impact users' acceptance toward mobile phone–based intervention [35,36], whether and how these factors influence the middle-agers' needs warrants research efforts. To summarize our research questions, this study aims to answer (1) how middle-agers perceive toward mobile phone–based exercise-promoting BCTs in terms of Kano quality categories; (2) whether current mobile phone apps meet middle-agers' expectations toward mobile phone–based exercise-promoting BCTs; and (3) whether and how the differences across levels of mobile phone self-efficacy (MPSE) and PA influence the quality perception.

Methods

Recruitment

This study was designed as a cross-sectional survey conducted from May to August 2015. We recruited participants from two sources. Most of the participants were recruited from those who volunteered to receive physical examinations performed by public health centers in Northern Taiwan. Posters describing study objectivities were displayed in venues where participants attended physical examinations. If an attendant approached recruitment posters, a research assistant introduced him or her to the survey and explained the aim of the study. To increase the number of participants, we also recruited participants in senior communities in Taiwan. All of the participants were told that participation was voluntary, and that all information disclosed would be confidential. All participants provided written consent before being involved in subsequent investigation in which they completed a self-administered

questionnaire. Each participant was thanked with a coupon worth of US $6.6.

Questionnaire Design

The questionnaire includes 4 sections regarding background, MPSE (10 items), PA (29 items), and quality perception toward BCTs via mobile phone (104 items). MPSE refers to mobile phone users' confidence to undertake specific tasks (eg, interacting with PA apps) [37]. This construct should be included in the research model, as eHealth studies are recommended to avoid generalizability issues arisen from assuming the reference population to be skilled in using a computer [35]. The 10-item Computer Self-Efficacy Scale was adapted for use in this study [38]. The items were modified to fit our research context of using mobile phone apps (eg, "if there was no one around to tell me what to do as I go, I could use mobile phone apps to manage PA" and "if I had never used a mobile phone app to manage PA before, I am confident of using such an app"). All items were scored on a 5-point Likert scale from "strongly disagree" to "strongly agree," with higher scores representing higher MPSE. We used tertiary split to divide this variable into three levels to examine the quality perception toward design features in each level of MPSE. The 10-item scale had high internal consistency (Cronbach alpha=.922).

Physical activity level may influence the effect of PA intervention [36], suggesting the necessity of incorporating this important variable. To measure PA level, the International Physical Activity Questionnaire (IPAQ) was developed to assess the frequency and duration of vigorous intensity, moderate intensity, and walking activity. This questionnaire has two versions available: the International Physical Activity Questionnaire-long form (IPAQ-LF) and the International Physical Activity Questionnaire-short form (IPAQ-SF). Several studies have suggested that the IPAQ has acceptable reliability and validity [39-42], including a 12-country study [43]. Since its invention, the IPAQ has become one of the most widely used PA questionnaires [41]. This scale, as a proxy for PA level, was used to examine the effect of technology-mediated intervention (eg, websites and email [44,45]), and to predict physical function (eg, pain facilitatory and inhibitory function [46]) and mental health (eg, risks of persistent late-life depression [47]).

This study adopted the long form version of IPAQ-Taiwan, modifying the original IPAQ-LF version with cultural adaptation. The IPAQ-Taiwan developers reported that the long form version had a content validity indice of .992, suggesting high language equivalence with the original English version. The consistency value for the English and Chinese versions in terms of intraclass correlation coefficients were .945, also indicating the appropriateness of reliability [42]. According to the usefulness guidelines suggested by Heesch et al [48], the 28-item IPAQ-Taiwan defines vigorous PA and moderate PA in "Introduction" section, followed by 5 sections requesting participants of their time spent in PA during the past 7 days in terms of different activity classes (eg, work-related and transport-related activity). The IPAQ-Taiwan clarifies activities mentioned as examples in the questionnaire, consistent with another suggestion by Heesch et al [48]. As with the original IPAQ, PA time in three levels (ie, vigorous-level, moderate-level, and walking) of four domains (ie, work, transport, domestic and garden, and leisure) was filled by participants. We followed the IPAQ group's scoring protocol to assess participants' level of PA, which can be found in Multimedia Appendix 1.

The questionnaire measuring quality perception toward BCTs via mobile phone was developed in this study. On the basis of the CALO-RE taxonomy, two of the authors developed 52 design features of mobile phone apps targeted at promoting PA, as illustrated in Table 2. If mobile phones can implement a BCT with more than one way, each way was designated as a design feature. For example, three design features (eg, contextual cues [A31], location cues [A32], and people cues [A33]) were derived from the BCT of using contextual cues.

For each design feature, a functional question asks participants' feelings in the case of fulfillment of the feature, and a dysfunctional question asks participants' feelings in the case of nonfulfillment of the feature [26,30,49]. Participants answered each question by choosing one of the five options: "Satisfied," "It should be that way," " I am indifferent," "I can live with it," and "Dissatisfied." For a specific participant, his or her quality perception toward the design feature can be determined by looking up the Kano evaluation matrix (Table 3) with the functional answer and the dysfunctional answer. There are five possible quality categories: *attractive* (A), *one-dimensional* (O), *must-be* (M), *reverse* (R), and *indifferent* (I) [49]. Assuming a nonlinear relationship between product performance and customer satisfaction [25,50,51], the Kano method defines an attribute is *attractive* (A) if an increase in the performance of an attribute enhances customer satisfaction, whereas a decrease in performance does not lead to dissatisfaction; an attribute is *one-dimensional* (O) if an increase in the performance of an attribute enhances customer satisfaction, whereas a decrease in performance also increases dissatisfaction; an attribute is *must-be* (M) if an increase in the performance does not increase satisfaction, but a deteriorating performance increases dissatisfaction; and an attribute is *indifferent* (I) if neither an increasing performance nor a decreasing performance can affect satisfaction. The definitions of reverse and questionable attributes can be found in [25,50]. The quality perception toward a design feature over a sample can be determined by selecting the highest frequency of quality categories for all the participants in the sample [49].

Table 2. A representative set of 52 design features based on the Coventry, Aberdeen, and London—Refined (CALO-RE) taxonomy.

Design features of exercise-promotion apps	Code
Apps provide information on consequences of exercise in general.	A1
Apps provide information on customized consequences of exercise.	A2
Apps provide information about others' approval of my exercise.	A3
Apps provide information about others' exercise status.	A4
Apps provide information about avoided movement in exercise.	A5
Apps help set exercise goals.	A6
Apps help set graded tasks by decomposing goals.	A7
Apps prompt review of exercise goals.	A8
Apps can check the extent to which previously set exercise goals were achieved.	A9
Apps remind me to record my exercise behavior.	A10
Apps can record my exercise behavior automatically.	A11
Apps can set the health goals to be achieved by exercise.	A12
Apps can prompt review of health goals.	A13
Apps can check the extent to which my expected goals were achieved.	A14
Apps remind me to keep records of my exercise outcome.	A15
Apps can automatically record my exercise outcome.	A16
Apps can assist me in detailed exercise planning.	A17
Apps can remind me to think about potential barriers in exercise planning.	A18
Apps can remind me to identify the ways of overcoming potential barriers when exercise planning.	A19
Apps prompt rewards contingent on effort toward exercise preparation.	A20
Apps provide rewards contingent on successful exercise.	A21
Apps provide graded use of contingent rewards over time.	A22
Apps prompt generalization of exercise.	A23
Apps remind me of past successful experience of exercise.	A24
Apps provide exercise records.	A25
Apps check the discrepancy between exercise performance and the set goals.	A26
Apps provide me with data about the discrepancy between my exercise performance and others'.	A27
Apps provide information on where and when to do the exercise.	A28
Apps provide instructions on how to do the exercise by text or voice.	A29
Apps show how to do the exercise through visual demonstrations.	A30
Apps can set context cues which remind me to exercise.	A31
Apps can set location cues which remind me to exercise.	A32
Apps can set people cues which remind me to exercise.	A33
Apps remind me to alter environment in ways so that it is more supportive of the exercise.	A34
Apps create the exercise goals as agreed behavioral contract.	A35
Apps prompt me to rehearse or repeat the exercise behavior numerous times.	A36
Exercise reminders are gradually reduced in intensity, duration, and frequency over time.	A37
Apps facilitate social comparison.	A38
Apps make it easy to elicit social support to my exercise from other people.	A39
Apps remind me to focus on partners who are the exercise role models.	A40
Apps facilitate the discussion with exercise role models.	A41
Apps induce perceptions of future regret about not doing exercise.	A42

Design features of exercise-promotion apps	Code
Apps provide risk information which evokes a fearful response.	A43
Apps prompt self-talk to encourage, support, and maintain exercise.	A44
Apps prompt mental imagery (to imagine initiating or maintaining exercise is easy).	A45
Apps provide strategies in advance to avoid sustainability problem of exercise.	A46
Apps provide stress management to reduce anxiety to facilitate the performance of the exercise.	A47
Apps remind me to attend motivational interviewing which can minimize resistance and resolve ambivalence to change.	A48
Apps assist time management to make time for exercise.	A49
Apps provide general communication skills training.	A50
Apps stimulate anticipation of future rewards.	A51
Apps can set exercise time reminders.	A52

Table 3. The Kano evaluation matrix.

Quality attribute	Dysfunctional answer					
Functional answer		Satisfied	It should be that way	I am indifferent	I can live with it	Dissatisfied
	Satisfied	Q[a]	A[b]	A	A	O[c]
	It should be that way	R[d]	I[e]	I	I	M[f]
	I am indifferent	R	I	I	I	M
	I can live with it	R	I	I	I	M
	Dissatisfied	R	R	R	R	Q

[a]Q: questionable.

[b]A: attractive.

[c]O: one-dimensional.

[d]R: reverse.

[e]I: indifferent.

[f]M: must-be.

Statistical Analysis

The analysis began with excluding invalid responses. As the chronological definition of middle age varies in existing studies (eg, 40-59 years [2], 45-74 years [52], and 40-60 years [53]), we included responses only from participants aged between 40 and 69 years. Next, missing values were identified, which was followed with a value substitution procedure. As the variables with more than one missing value were all in the Kano questionnaire, the importance of the Kano variables suggested the use of multiple imputations. We opted to include all the nonmissing variables as predictors in the initial imputation model. With multinomial logistic regression, the variables with significant predictability were identified. As suggested in [54], we used three imputations for every missing value. As the Kano method determines the categories with the largest number of votes, we created the final set of winning categories by forming the union of winning categories in three imputations. The analysis revealed that the winning categories in each imputation were the same, which may be attributed to the small proportions of missing values. After the multiple imputation procedure, all the data entered the Kano analysis.

The Kano Analysis

The Kano method determines the highest frequency of quality categories in a design feature as the winning category, as these categories represent the dominant customer view [55]. Previous study applying the Kano method suggested the use of two additional measures, category strength, and total strength, to determine whether there exists more than one attribute that dominates [56]. Category strength is defined as “the extent of how firmly the participants felt that an item was in one category or another” [56]. We calculated the difference (in percentage) between the highest and the second highest categories to measure category strength. A category strength value greater than 6% indicates a statistical difference between the highest and the second highest categories [56]. Total strength is defined as the total proportion of positive attributes (ie, *attractive* [A], *one-dimensional* [O], and *must-be* [M]). According to [56], if the category strength of an attribute were lower than 6% and the total strength exceeded 60%, then it could be statistically impossible to classify the attribute in one category or another (referred to as the Lee and Newcomb rule hereafter). Because 60% could be determined arbitrarily, such an attribute would fall into the mixed (X) category [55], indicating that a design feature may turn out to be determined as multiple categories. We adopted an aggressive position in which design features

deserve recommendations as long as the categorization results include any positive category. This position emphasizes the importance of customer demands while also creating opportunities of adding value to mobile phone–based exercise promotion.

Two-Factor Analysis

To realize the predictability of two independent variables, we created a multinomial regression model for each of those design features. By regressing quality perception on PA level and MPSE, the fit of the models was examined and the likelihood ratio tests were conducted to ascertain the significance of predictors.

Institutional Review Board (IRB)

Ethical approval was granted by the Institutional Review Board (IRB) and Chang Gung Memorial Hospital (CGMH) (103-2125B and 104-3029C). Permission for data collection was also obtained from the officials of the public health centers. The participants were informed about the study, its importance, and confidentiality of the information collected. They were also told that they owned rights to leave the study at any time before signing their written consent for participation in the study. All participants' data are maintained in a secure manner by separating participants' identifiers and associated data, as recommended by the IRB.

Results

Sample Demographics

The sample size was 105. The participants had an average age of 55.7 years. The sex ratio of participants in the nonmissing responses was 53:50. Most participants (73/105; 69.5%) had at least a senior high school education and more than a half in the nonmissing responses (53/98; 54%) had an annual income of NT $720,000 (US $23,630) or more. More than half of the participants (n=58) used mobile phone apps longer than half an hour per day, and about one-fifth (n=20) played apps longer than 2 h. Demographic data are presented in Table 4. We also asked the participants about their most frequently used mobile phone apps. The result showed that the LINE (n=62) and Facebook apps (n=11) were respectively ranked in the first and second places, suggesting the popularity of social networking apps.

Kano Analysis

This section proposes the analysis results obtained in conducting Kano analysis on the total sample and the subsamples by PA and by MPSE. When leading categories had close votes, we used the Lee and Newcomb rule to determine whether to list multiple winners. When winning categories of a design feature included the *indifferent* category and at least one positive category (ie, *attractive*, *one-dimensional*, *must-be*), only positive categories were described in-text, as positive categories are more informative to app developers. Nevertheless, we reported in Tables 5-10 with an expression X(P, I) to indicate a tie between a positive category P and the *indifferent* category (I). To make this paper concise, design features are not shown in Tables 5-10 unless their categorization results include at least one positive category, whereas complete results can be found in Multimedia Appendix 2.

Table 5 reports the categorization results based on the responses from all participants who filled out the Kano questionnaire (n=103). It was found that 51 of the 52 design features were classified as *indifferent*, suggesting that these 51 design features did not interest the subjects in the total sample. One design feature (A30: visual demonstrations) may be categorized as *attractive (28)*, despite with a tie with the *indifferent* category.

The categorization results in the subsamples of PA levels were based on 102 valid responses from those participants who completed both the Kano questionnaire and the IPAQ. We first examined the high-PA participants and found that all of the 52 design features were categorized as *indifferent*. As no design features with a smaller-than 0.06 category strength had a total strength larger than 60%, we had sufficient confidence that these designs did not motivate this subsample.

Table 6 shows the categorization results of the medium-PA participants. For this specific subsample, mobile phone apps offered limited incentives to use. It was found that 46 of the 52 design features were determined as one category of *indifferent*. A10 (reminding to record PA) was categorized as *must-be (7)*, suggesting that the participants with medium-PA considered as granted the design feature of reminding to record exercise behavior. Five other design features had close proportions in the *indifferent* category and a positive category, including A5 (offering movements to be avoided) categorized into *one-dimensional (5)*; A11 (automatically record physical activity [PA]) into *attractive (6)*; A20 (contingent reward for exercise preparation) into *must-be (5)*; A23 (prompt generalization of exercise) into *one-dimensional (6)*; and A29 (exercise instruction with text or voice) into *one-dimensional (6)*. Most of the functions needed are related to information provision and behavioral facilitation.

Table 4. Descriptive statistics on sample demographics (N=105).

Variable		n (%)
Gender		
	Male	53 (50.5)
	Female	50 (47.6)
	Missing	2 (1.9)
Education		
	Junior high school or less	30 (28.6)
	Senior high school	33 (31.4)
	Bachelor's degree	39 (37.1)
	Graduate degree	1 (1.0)
	Missing	2 (1.9)
Employment		
	Employed	74 (70.5)
	Unemployed or retired	29 (27.6)
	Missing	2 (1.9)
Marital status[a]		
	Married	84 (80.0)
	Widowed	7 (6.7)
	Divorced	8 (7.6)
	Not married	3 (2.9)
	Missing	3 (2.9)
Monthly household income[a]		
	≤NT $39,999 (US $1313)	22 (21.0)
	≤NT $49,999 (US $1641)	12 (11.4)
	≤NT $59,999 (US $1969)	11 (10.5)
	≤NT $69,999 (US $2297)	9 (8.6)
	≤$NT $79,999 (US $2626)	8 (7.6)
	≤NT $89,999 (US $2954)	12 (11.4)
	>NT $90,000 (US $2955)	24 (22.9)
	Missing	7 (6.7)
Daily app using time		
	≤30 min	43 (41.0)
	≤120 min	38 (36.2)
	≤240 min	16 (15.2)
	>240 min	4 (3.8)
	Missing	4 (3.8)

[a]Percentages may not add up to 100 due to rounding.

Table 5. Categorizing design features by the total sample (n=103).

Design features	Frequency of design feature						Category strength (%)	Total strength (%)	Classification results
	A	M	O	I	R	Q			
A30	28	15	23	33	1	3	5	64	X(I, A)[a]

[a]X(C_1, C_2) indicates that a design feature had close proportions in two categories of C_1 and C_2.

Table 6. Categorizing design features by medium physical activity participants (n=18).

Design features	Frequency of design feature						Category strength (%)	Total strength (%)	Classification results
	A	M	O	I	R	Q			
A5	4	3	5	6	0	0	6	67	X(I, O)[a]
A10	4	7	2	5	0	0	11	72	M
A11	6	5	1	6	0	0	0	67	X(A, I)[a]
A20	4	5	3	6	0	0	6	67	X(I, M)[a]
A23	2	3	6	7	0	0	6	61	X(I, O)[a]
A29	4	1	6	7	0	0	6	61	X(I, O)[a]

[a]X(C_1, C_2) indicates that a design feature had close proportions in two categories of C_1 and C_2.

In contrast, low-PA users need much more support from mobile phone–based apps. Table 7 shows the categorization results of low-PA participants. In total, 31 design features were categorized only as *indifferent*. It was found that 6 design features (A1: general consequences of exercise; A2: customized consequences of exercise; A13: browse health goals; A22: contingent rewards with grading; A25: PA history; and A26: comparing actual PA with PA goal) were categorized as *must-be*, and 1 design feature (A14: compare actual health outcomes with health goals) was categorized into *one-dimensional*. The remaining 14 design features were categorized into more than one category according to the Lee and Newcomb rule. Among these 14 design features, 7 were classified into only positive categories, including A7 (help goal decomposition: *one-dimensional* (7) and *must-be* (6)), A10 (remind to record PA: *one-dimensional* (7) and *must-be* (6)), A15 (remind to record health outcomes: *one-dimensional* (7) and *must-be* (7)), A21 (contingent reward for exercise practice: *one-dimensional* (7) and *must-be* (7)), A23 (prompt generalization of exercise: *one-dimensional* (7) and *must-be* (6)), A29 (exercise instruction with text or voice: *attractive* (7) and *must-be* (7)), and A30 (visual demonstration: *attractive* (7) and *must-be* (7)). The other 7 design features were determined as multiple categories including *indifferent*: A6 (set PA goals: *must-be* (7)), A8 (browse PA goals: *must-be* (7)), A9 (check goal conversions: *must-be* (7)), A12 (set health goals: attractive (5), *must-be* (6), *one-dimensional* (5)), A20 (contingent reward for exercise preparation: *one-dimensional* (6)), A24 (remind past success: *attractive* (6) and *must-be* (6)), and A52 (reminding to PA: *attractive* (6), *must-be* (7)). Low-PA users required assistance in goal management and time management. Furthermore, 6 motivational techniques (ie, A1, A2, A20, A21, A22, and A24) were considered either as *must-be* or as *one-dimensional*, suggesting the importance of extrinsic motivation to low-PA users.

A further analysis on the 21 design features revealed that only two (ie, A14 and A20) were not categorized as *must-be*, and all the 5 design features categorized into *attractive* were also categorized as *must-be*. Therefore, this finding suggested that these designs may be more of necessity than of attractiveness to low-PA participants.

The categorization results across levels of MPSE were based on 102 valid responses from those participants who completed both the Kano questionnaire and the MPSE questionnaire. Table 8 presents the categorization results of participants with high self-efficacy. A total of 4 design features (A23: prompt generalization of exercise; A26: comparing actual PA with PA goal; A29: exercise instruction with text or voice; and A30: visual demonstrations) were categorized as *one-dimensional*. It was found that 10 design features were categorized into more than one category including *indifferent*, including A2 (customized consequences of exercise: *must-be* (10)), A10 (remind to record PA: *must-be* (10)), A12 (set health goals: *one-dimensional* (11)), A13 (browse health goals: *one-dimensional* (10)), A14 (compare actual health outcomes with health goals: *one-dimensional* (11)), A15 (remind to record health outcomes: *one-dimensional* (10)), A21 (contingent reward for exercise practice: *one-dimensional* (10)), A22: (contingent rewards with grading: *one-dimensional* (14)); A25 (PA history: *one-dimensional* (12)), and A52 (reminding to PA: *one-dimensional* (10)). A breakdown analysis on the 14 design features revealed that, except with 2 design features (ie, A2 and A10) categorized as *must-be*, the other 12 design features were all *one-dimensional*. Therefore, this analysis suggested that users with high MPSE would consider these designs as more satisfactory as the mobile phone apps perform better in terms of these design features. The remaining 38 design features were all categorized as *indifferent*.

Table 7. Categorizing design features by low physical activity participants (n=22).

Design features	Frequency of design feature						Category strength (%)	Total strength (%)	Classification results
	A	M	O	I	R	Q			
A1	1	11	2	8	0	0	14	64	M
A2	1	8	4	8	0	1	0	59	M
A6	2	7	5	8	0	0	5	64	X(I, M)[a]
A7	3	6	7	5	1	0	5	73	X(O, M)[a]
A8	4	7	4	7	0	0	0	68	X(M, I)[a]
A9	3	7	4	8	0	0	5	64	X(I, M)[a]
A10	3	6	7	5	0	1	5	73	X(O, M)[a]
A12	5	6	5	5	1	0	5	73	X(M, A, O, I)[a]
A13	3	8	5	6	0	0	9	73	M
A14	3	6	8	3	1	1	9	77	O
A15	4	7	7	3	0	1	0	82	X(M, O)[a]
A20	4	5	6	7	0	0	5	68	X(I, O)[a]
A21	4	7	7	3	0	1	0	82	X(M, O)[a]
A22	5	8	6	3	0	0	9	86	M
A23	4	6	7	4	1	0	5	77	X(O, M)[a]
A24	6	6	4	6	0	0	0	73	X(A, M, I)[a]
A25	4	8	5	5	0	0	14	77	M
A26	4	9	5	4	0	0	18	82	M
A29	7	7	5	3	0	0	0	86	X(A, M)[a]
A30	7	7	5	3	0	0	0	86	X(A, M)[a]
A52	6	7	3	6	0	0	5	73	X(M, A, I)[a]

[a]$X(C_1, C_2, ..., C_n)$ indicates that a design feature had close proportions in the categories of C_i, $1 \leq i \leq n$.

Table 9 presents the categorization results of medium MPSE participants. A total of 46 design features were categorized as *indifferent* and 4 design features (ie, A15: remind to record health outcomes; A22: contingent rewards with grading; A25: PA history; and A52: reminding to PA) were categorized as *must-be*. A23 (prompt generalization of exercise) and A26 (comparing actual PA with PA goal) were determined as a tie with two categories of *indifferent* and *must-be* (ie, A23: *must-be* (10); A26: *must-be* (13)). This analysis suggested that this specific subsample had weak intention to use mobile phone apps to promote exercise and tended to consider the functions as of necessity.

Table 8. Categorizing design features by high self-efficacy participants (n=35).

Design features	Frequency of design feature						Category strength (%)	Total strength (%)	Classification results
	A	M	O	I	R	Q			
A2	6	10	7	11	0	1	3	66	X(I, M)[a]
A10	6	10	8	10	0	1	0	69	X(M, I)[a]
A12	6	5	11	12	0	1	3	63	X(I, O)[a]
A13	7	7	10	9	0	2	3	69	X(O, I)[a]
A14	6	6	11	10	0	2	3	66	X(O, I)[a]
A15	5	8	10	10	0	2	0	66	X(O, I)[a]
A21	7	5	10	11	0	2	3	63	X(I, O)[a]
A22	4	3	14	12	1	1	6	60	X(O, I)[a]
A23	6	3	14	10	0	2	11	66	O
A25	6	4	12	11	0	2	3	63	X(O, I)[a]
A26	5	6	13	9	0	2	11	69	O
A29	7	3	15	8	0	2	20	71	O
A30	9	3	13	7	1	2	11	71	O
A52	6	6	10	12	0	1	6	63	X(I, O)[a]

[a]X(C_1, C_2) indicates that a design feature had close proportions in two categories of C_1 and C_2.

Table 9. Categorizing design features by medium self-efficacy participants (n=32).

Design features	Frequency of design feature						Category strength (%)	Total strength (%)	Classification results
	A	M	O	I	R	Q			
A15	3	12	4	12	0	1	0	59	M
A22	7	12	5	8	0	0	13	75	M
A23	6	10	4	10	2	0	0	63	X(M,I)[a]
A25	6	13	3	10	0	0	9	69	M
A26	4	13	3	12	0	0	3	63	X(M,I)[a]
A52	7	12	3	10	0	0	6	69	M

[a]X(C_1, C_2) indicates that a design feature had close proportions in two categories of C_1 and C_2.

Table 10. Categorizing design features by low self-efficacy participants (n=35).

Design features	Frequency of design feature						Category strength (%)	Total strength (%)	Classification results
	A	M	O	I	R	Q			
A30	12	4	7	12	0	0	0	66	X(A, I)[a]

[a]X(C_1, C_2) indicates that a design feature had close proportions in two categories of C_1 and C_2.

Table 10 presents the categorization results by the low MPSE subsample. Only A30 (visual demonstration) was categorized as both *attractive* and *indifferent*. All of the remaining designs were *indifferent*. This subsample exhibited low interest in those design features to increase PA.

Next, we compared users' demands with mobile phone–based BCTs and the supply in the mobile phone apps market. Table 11 compares the categorization results across the total sample and subsamples. If the quality categories in a cell are shown with a superscript *i*, this means that these categories are tied

with the *indifferent* category. According to [34], the rightmost field in Table 11 provides the prevalence of particular BCTs delivered by popular mobile phone apps. The range of prevalence numbers of demanded design features was [0%, 49%] with an average of 20% (4.3/21). The two top-ranked features (ie, social support [A39, 79%] and social approval [A3, 64%], also listed in Table 11) were available in more than half of the apps inspected. However, both features were considered as *indifferent* by our sample and in all of the subsamples, either by PA level or by MPSE. Exercise instruction (A29, 49%) and visual demonstration (A30, 47%), ranked in the next two places, were considered as needed. On the other hand, design features that can contribute to fulfill users' needs across different user groups are available in less than 40% of the apps inspected. For example, prompting exercise generalization (A23, 0%) is valuable to four subsamples of middle-agers, but was not found in the inspected apps. Prompts or cues (A52) can create value to three subsamples (ie, low-PA, high-MPSE, and medium-MPSE), but only 35% of the apps offered this feature. Table 11 also showed that reminding to record PA (A10, 29%), reminding to record health outcomes (A15, 22%), and contingent rewards with grading or shaping (A22, 1%) were unavailable in most of the apps but each feature was needed in three subsamples. Furthermore, 5 design features were needed by two subsamples but the prevalence numbers were not larger than 30% (eg, browse health goals [A13, 6%] and contingent reward for exercise practice [A21, 3%]). These findings suggest that current supply of exercise-promoting features in mobile phone apps might not fully match middle-aged adults' needs. The gaps represent strategic opportunities for app designers to fulfill the needs of customers with customized characteristics.

Two-Factor Analysis

To realize the effects of PA level and MPSE on quality categorization (ie, the third research question), we created one multinomial logistic regression model for each design feature. To avoid zero frequency, we combined the *attractive* and *one-dimensional* categories into a new *valued* category. The *must-be* category remained intact, whereas all of the other instances were entered into the *indifferent* category. The left-hand part of Multimedia Appendix 3 provides the Pearson and deviance statistics for model fitting. As the deviance statistics of three models (ie, A20, A32, and A43) were significant, which suggested significant deviation from observations to predicted values, we excluded these models from the likelihood ratio tests that followed.

The remaining 49 models were examined with the likelihood ratio tests to ascertain the significance of predictors to the models. The chi-square statistics and the significance of coefficients for two predictor variables (ie, MPSE and PA level) were listed in the right-hand part of the Multimedia Appendix 2. MPSE was significant in 14 models (ie, A3, A6, A8, A10, A13, A18, A22, A25, A26, A29, A35, A41, A45, and A48), whereas only 1 (ie, A3) was found as significant for PA level. This analysis revealed that combining MPSE and PA could significantly predict quality categorization on A3, whereas only MPSE was a significant predictor predicting other 13 design features.

To delve into the effect of different levels of predictors, we analyzed the effects of coefficients on design satisfaction with regard to each of the 14 design features. We first analyzed the effect of PA on categorizing A3 (social approval). With the medium-PA group as the baseline, neither the estimates for high-PA (b=−0.550, Wald χ^2_1=0.5 odds ratio, OR 0.58 [95% CI 0.12-2.82], P=.50) nor the estimates for low-PA (b=−1.453, Wald χ^2_1=1.4; OR 0.23 [95% CI 0.02-2.53], P=.24) for the *must-be* category (compared with the *indifferent* category) was significant. For the *valued* category (compared with the *indifferent* category), the estimates for high-PA (b=19.703, Wald χ^2_1=540.5; OR 3.60E-8 [95% CI 6.85E-7 to 1.90E-9], P<.001) was significant. This suggested that high-PA participants were more likely to categorize A3 into *attractive* or *one-dimensional* than the medium-PA group.

We next turned to assess the estimates of MPSE coefficients. Table 12 presents the b values, the Wald statistics, the values of significance of testing the estimates across the MPSE levels, and 95% CI for OR for the *must-be* category compared with the *indifferent* category. The analysis revealed that, compared with their low MPSE counterparts, medium MPSE participants more likely categorized 5 design features as *must-be* than as *indifferent*. More specifically, this medium MPSE group considered the following features as *must-be*: A6 (set PA goals: b=−1.712, Wald χ^2_1=5.4; OR 0.18 [95% CI 0.04-0.77], P=.02), A22 (contingent rewards with grading: b=−1.790, Wald χ^2_1=5.7; OR 0.17 [95% CI 0.04-0.72], P=.02), A25 (PA history: b=−1.590, Wald χ^2_1=5.5; OR 0.20 [95% CI 0.05-0.77], P=.02), A26 (comparing actual PA with PA goal: b=−1.370, Wald χ^2_1=4.2; OR 0.25 [95% CI 0.07-0.94], P=.04), and A45 (prompt mental imagery: b=−2.770, Wald χ^2_1=6.1; OR 0.06 [95% CI 0.01-0.56], P=.01). Besides, compared with high-MPSE participants, medium MPSE participants more likely considered A22 (contingent rewards with grading) as *must-be* than as *indifferent* (b=−1.932, Wald χ^2_1=5.5; OR 0.15 [95% CI 0.03-0.73], P=.02), as illustrated in Table 13. The OR indicates that the change in the odds of categorizing A22 as *must-be* compared with *indifferent* is .15 as the MPSE level changes from medium to high, which suggests that this design feature was more likely categorized to *must-be* (compared with *indifferent*) by medium MPSE participants than by high-MPSE participants. As Table 14 indicates that no design feature was found with a significant regression coefficient, this analysis found no evidence to argue that the categorization likelihood differs between the high-MPSE and low-MPSE subsamples.

Table 11. Positive quality categories by design features and participant characteristics.

Design feature	Total sample	PA[a] (H)	PA (M)	PA (L)	MPSE[b] (H)	MPSE (M)	MPSE (L)	Prevalence (%) [34]
A1 (general consequences of exercise)[c]				M				2
A2 (customized consequences of exercise)[c]				M	M[g]			2
A3 (social approval)								64
A5 (offering movements to be avoided)[f]			O[g]					N/A[h]
A6 (set PA goals)				M[g]				36
A7 (help goal decomposition)				(O, M)				33
A8 (browse PA goals)				M[g]				17
A9 (check goal conversions)				M[g]				8
A10 (remind to record PA)			M	(O, M)	M[g]			29
A11 (automatically record PA)			A[g]					29
A12 (set health goals)				(A, O, M)[g]	O[g]			17
A13 (browse health goals)[d]				M	O[g]			6
A14 (compare actual health outcomes with health goals)[d]				O	O[g]			6
A15 (remind to record health outcomes)				(O, M)	O[g]	M		22
A20 (contingent reward for exercise preparation)[f]			M[g]	O[g]				N/A
A21 (contingent reward for exercise practice)				(O, M)	O[g]			3
A22 (contingent rewards with grading or shaping)[e]				M	O[g]	M		1
A23 (prompting exercise generalization)[f]			O[g]	(O, M)	O	M[g]		0
A24 (remind past success)				(A, M)[g]				4
A25 (PA history)				M	O[g]	M		42
A26 (comparing actual PA with PA goal)				M	O	M[g]		42
A29 (PA instruction with text or voice)			O[g]	(A, M)	O			49
A30 (visual demonstration)	A[g]			(A, M)	O		A[g]	47
A39 (social support)								79
A52 (reminding to PA)				(A, M)[g]	O[g]	M		35
Number of features	1	0	6	21	14	6	1	

[a]PA: physical activity.

[b]MPSE: mobile phone self-efficacy.

[c]Corresponding to information about health consequences in [34].

[d]Corresponding to review outcome goals in [34].

[e]Corresponding to reward approximation in [34].

[f]No observation reported in [34].

[g]Tied with the *indifferent* category.

[h]N/A: not applicable.

Table 12. Parameter estimates of change of mobile phone self-efficacy (MPSE) from medium to low (*must-be* over *indifferent*).

Design feature	Mobile phone self-efficacy (H vs L)				95% CI for odds ratio (lower bound-upper bound)
	b	Wald	Significance	Odds ratio	
A3	−0.592	0.4	.54	0.55	0.08-3.70
A6	−1.712	5.4	.02	0.18	0.04-0.77
A8	−0.778	1.2	.27	0.46	0.12-1.84
A10	−0.673	1.1	.30	0.51	0.14-1.81
A13	−0.995	2.3	.13	0.37	0.10-1.34
A18	−0.551	0.5	.50	0.58	0.12-2.88
A22	−1.790	5.7	.02	0.17	0.04-0.72
A25	−1.590	5.5	.02	0.20	0.05-0.77
A26	−1.370	4.2	.04	0.25	0.07-0.94
A29	−0.981	1.8	.18	0.38	0.09-1.58
A35	−0.655	0.3	.61	0.52	0.04-6.21
A41	−1.311	1.2	.28	0.27	0.03-2.86
A45	−2.770	6.1	.01	0.06	0.01-0.56
A48	−1.533	3.1	.08	0.22	0.04-1.19

Table 13. Parameter estimates of change of mobile phone self-efficacy (MPSE) from medium to high (*must-be* over *indifferent*).

Design feature	Mobile phone self-efficacy (H vs L)				95% CI for odds ratio (lower bound-upper bound)
	b	Wald	Significance	Odds ratio	
A3	0.494	0.4	.55	1.64	0.32-8.38
A6	−0.438	0.5	.48	0.65	0.19-2.16
A8	−0.144	0.0	.83	0.87	0.24-3.12
A10	0.210	0.1	.73	1.23	0.38-4.03
A13	−0.119	0.0	.85	0.89	0.25-3.13
A18	−20.685	N/A[a]	N/A	1.04E-9	1.04E-9 to 1.04E-9
A22	−1.932	5.5	.02	0.15	0.03-0.73
A25	−1.389	3.6	.06	0.25	0.06-1.05
A26	−0.558	0.7	.41	0.57	0.15-2.16
A29	−0.841	1.1	.31	0.43	0.09-2.15
A35	0.224	0.0	.83	1.25	0.16-9.71
A41	−19.646	N/A	N/A	2.94E-9	2.94E-9 to 2.94E-9
A45	−1.010	2.1	.15	0.36	0.09-1.45
A48	−21.029	N/A	N/A	7.37E-10	7.37E-10 to 7.37E-10

[a]N/A: not applicable.

Table 14. Parameter estimates of change of mobile phone self-efficacy (MPSE) from low to high (*must-be* over *indifferent*).

Design feature	Mobile phone self-efficacy (H vs L)				95% CI for odds ratio (lower bound-upper bound)
	b	Wald	Significance	Odds ratio	
A3	1.085	1.4	.24	2.96	0.48-18.11
A6	1.274	2.7	.10	3.58	0.78-16.73
A8	0.634	0.7	.41	1.89	0.42-8.37
A10	0.882	1.7	.19	2.42	0.64-9.06
A13	0.876	1.5	.22	2.40	0.59-9.78
A18	−20.134	N/A[a]	N/A	1.80E-9	0.00-0.00
A22	−0.142	0.0	.87	0.87	0.16-4.85
A25	0.201	0.1	.80	1.22	0.27-5.64
A26	0.812	1.2	.27	2.25	0.53-9.56
A29	0.140	0.0	.88	1.15	0.20-6.61
A35	0.880	0.5	.49	2.41	0.20-28.68
A41	−18.335	N/A	N/A	1.09E-8	0.00-0.00
A45	1.761	2.3	.13	5.82	0.59-57.22
A48	−19.496	N/A	N/A	3.41E-9	0.00-0.00

[a]N/A: not applicable.

We next turned to analyze the effect of MPSE on the relative likelihoods of the *valued* category versus the *indifferent* category, as illustrated in Tables 15-17. In comparing the high group to the medium group, MPSE was able to predict the relative likelihoods with regard to 5 design features, as indicated in Table 15. Specifically, the high-group is more likely considered as *valued* in A3 (social approval: *b*=2.622, Wald χ^2_1=5.6; OR 13.77 [95% CI 1.57-120.51], *P*=.02), A8 (browse PA goals: *b*=1.540, Wald χ^2_1=5.8; OR 4.66 [95% CI 1.33-16.37], *P*=.02), A10 (reminding to record PA: *b*=1.522, Wald χ^2_1=5.0; OR 4.58 [95% CI 1.21-17.39], *P*=.03), A13 (browse health goals: *b*=1.673, Wald χ^2_1=6.4; OR 5.33 [95% CI 1.47-19.37], *P*=.01), A35 (PA fulfilled as a contract: *b*=20.051, Wald χ^2_1=719.9; OR 5.11E-8 [95% CI 1.18E-8 to 2.21E-9], *P*<.001), and A41 (Talk to role models: *b*=1.573, Wald χ^2_1=4.9; OR 4.82 [95% CI 1.19-19.49], *P*=.03). The high-group also categorized A26 as *valued*, compared with the low-group, as suggested by the regression coefficient and the Wald statistic (comparing actual PA with PA goal: *b*=1.237, Wald χ^2_1=4.8; OR 3.444 [95% CI 1.15-10.35], *P*=.03) as shown in Table 16. Moreover, the low group, compared with their medium counterpart, had a relative advantage in categorizing A10 (remind to record PA) into *valued* rather than *indifferent* (*b*=1.324, Wald χ^2_1=4.0; OR 3.76 [95% CI 1.02-13.85], *P*=.047). The OR indicates that the change in the odds of categorizing A10 as *valued* compared with *indifferent* is 3.76 as the MPSE level changes from medium to low, which is shown in Table 17.

For the significant relationships shown in Tables 12-17,Table 18 summarizes the design features and quality categories (in parenthesis) with the comparing MPSE level listed in the column header and the reference MPSE level in the row header. For example, A8 (browse PA goals; A/O) is listed under the column H (high) and the row M (medium) because high-MPSE participants, compared with medium-MPSE participants, are more likely to categorize this design feature into *attractive* or *one-dimensional*. This table indicates that high-MPSE participants more likely categorize design features into the *valued* category than the other two groups, as *6* design features were associated with relative likelihoods of *valued* versus *indifferent* by the high-MPSE group, whereas only *1* design feature was associated by the low group and *none* by the medium group in contrast. The finding also revealed that the medium group more likely categorizes design features into *must-be*, as all of the six relationships categorized into the *must-be* category were found when medium-MPSE was compared with the other two MPSE levels.

The third research question can be answered with integrating the analysis results from the Kano analysis and the multinomial regression. The two predictors could influence customers' quality perception to a varying degree. As the Kano analysis revealed an increasing trend in the number of positive design features as PA decreases, PA might have a negative influence. However, the multinomial regression suggested changes of PA only influenced quality categorization in A3. Since A3 was not the winning category as indicated in Table 11, the effect by PA level could be very small. In contrast, MPSE was tested as significant in 14 regression models, of which 6 were the dominant categories according to Table 18. The reasoning suggests that the influence of MPSE on quality perception should be stronger than that of PA.

Table 15. Parameter estimates of change of mobile phone self-efficacy (MPSE) from medium to high (*valued* [*attractive+one-dimensional*] over *indifferent*).

Design feature	Mobile phone self-efficacy (H vs L)				95% CI for odds ratio (lower bound-upper bound)
	b	Wald	Significance	Odds ratio	
A3	2.622	5.6	.02	13.77	1.57-120.51
A6	0.930	2.2	.14	2.54	0.75-8.61
A8	1.540	5.8	.02	4.66	1.33-16.37
A10	1.522	5.0	.03	4.58	1.21-17.39
A13	1.673	6.4	.01	5.33	1.47-19.37
A18	0.788	1.8	.18	2.20	0.70-6.91
A22	-0.120	0.0	.84	0.89	0.28-2.86
A25	0.511	0.7	.39	1.67	0.52-5.38
A26	1.154	3.4	.07	3.17	0.92-10.88
A29	1.112	3.4	.06	3.04	0.94-9.83
A35	20.051	719.9	.00	5.11E-8	1.18E-8 to 2.21E-9
A41	1.573	4.9	.03	4.82	1.19-19.49
A45	0.431	0.5	.47	1.54	0.48-4.92
A48	0.692	1.2	.27	2.00	0.59-6.83

Table 16. Parameter estimates of change of mobile phone self-efficacy (MPSE) from low to high (*valued* [*attractive + one-dimensional*] over *indifferent*).

Design feature	Mobile phone self-efficacy (H vs L)				95% CI for odds ratio (lower bound-upper bound)
	b	Wald	Significance	Odds ratio	
A3	0.948	2.1	.15	2.58	0.71-9.43
A6	0.698	1.7	.20	2.01	0.70-5.78
A8	0.309	0.3	.56	1.36	0.48-3.84
A10	0.198	0.1	.73	1.22	0.40-3.73
A13	0.887	2.6	.11	2.43	0.83-7.13
A18	0.189	0.1	.72	1.21	0.43-3.38
A22	0.569	1.2	.28	1.77	0.63-4.97
A25	1.048	3.7	.05	2.85	0.98-8.26
A26	1.237	4.8	.03	3.44	1.15-10.35
A29	0.908	2.8	.10	2.48	0.85-7.25
A35	1.010	1.8	.18	2.75	0.64-11.88
A41	0.398	0.6	.46	1.49	0.52-4.24
A45	0.455	0.7	.40	1.58	0.55-4.52
A48	0.213	0.2	.69	1.24	0.43-3.57

Table 17. Parameter estimates of change of Mobile phone self-efficacy (MPSE) from low to medium (*valued* [*attractive* + *one-dimensional*] over *indifferent*).

Design feature	Mobile phone self-efficacy (H vs L)				95% CI for odds ratio (lower bound-upper bound)
	b	Wald	Significance	Odds ratio	
A3	1.674	2.2	.14	5.33	0.57-49.82
A6	0.232	0.1	.71	1.26	0.37-4.25
A8	1.231	3.8	.05	3.42	0.99-11.83
A10	1.324	4.0	.047	3.76	1.02-13.85
A13	0.786	1.5	.23	2.20	0.62-7.81
A18	0.599	1.0	.32	1.82	0.56-5.93
A22	−0.689	1.3	.25	0.50	0.16-1.62
A25	−0.537	0.8	.38	0.59	0.18-1.93
A26	−0.083	0.0	.90	0.92	0.27-3.15
A29	0.203	0.1	.72	1.23	0.40-3.79
A35	19.041	N/A[a]	N/A	1.86E-8	1.86E-8 to 1.86E-8
A41	1.176	2.6	.11	3.24	0.77-13.69
A45	−0.024	0.0	.97	0.98	0.298-3.194
A48	0.480	0.5	.46	1.62	0.453-5.764

[a]N/A: not applicable.

Table 18. The quality categories of design features with significant mobile phone self-efficacy (MPSE) coefficients.

	H[a]	M[a]	L[a]
H[b]		A22: contingent rewards with grading or shaping (M)[c]	
M[b]	A3: social approval (A/O) A8: browse PA[d] goals (A/O) A10: remind to record PA (A/O) A13: browse health goals (A/O)[e] A35: PA fulfilled as a contract (A/O) A41: talk to role models (A/O)		A10: remind to record PA (A/O)
L[b]	A26: comparing actual PA with PA goal (A/O)[e]	A6: set PA goals (M) A22: contingent rewards with grading or shaping (M)[c] A25: PA history (M)[c] A26: comparing actual PA with PA goal (M)[c] A45: prompt mental imagery (M)	

[a]Comparing level represented by a dummy variable.

[b]Reference level.

[c]This category also listed as a winning category in the Kano analysis on the medium-PA participants.

[d]PA: physical activity.

[e]This category also listed as a winning category in the Kano analysis on the high-PA participants.

Discussion

Principal Findings

The Kano analyses provide evidence with which to answer the research questions. Overall, BCT-based exercise-promoting features that can attract middle-agers are limited. The analysis on the total sample revealed that visual demonstration of exercise instructions (A30) may be the only *attractive* design feature, whereas the other 51 design features are perceived as *indifferent*. This result is not surprising, as studies have reported physical inactivity in the middle-aged population [57,58], explaining a potential lack of motivation to use these persuasive features. Our Kano analyses also suggest 6 positive design

features for mobile phone users with middle-PA, 21 for low-PA users, whereas also recommending 14 for high-MPSE, 6 for medium-MPSE, and 1 for low-MPSE users. The second research problem is answered by comparing the Kano analysis results to the prevalence numbers proposed in [34]. The analysis suggests that the implementation of demanded features could be low, as the average prevalence of demanded design features is 20% (4.3/21).

The third research question is answered with the findings obtained by the Kano analysis and the multinomial regression analyses. We found that both PA and MPSE could influence customers' quality perception, whereas the magnitude of effect is larger for MPSE, because MPSE effects quality perception of more design features than PA.

Implications for Design Features

Customization has been proposed as important to mHealth apps [12,59], assuming that users with different characteristics are associated with different needs. Our finding is consistent with this assumption, as the quality perception differs across the levels of PA and MPSE. Accordingly, apps should measure users' PA and MPSE for customization settings. However, inappropriate customization (eg, too many or incorrect features) will overload users with cognitive complexity which increases errors and reduces operational efficiency. Besides, adding more choices and options to a single user interface will create uncertainty and induce distraction, finally leading to negative experience [60]. Therefore, customization requires knowledge regarding the right functions delivered to the right users who need them. This study contributes with the design recommendations grounded in the Kano study with which exercise-promotion apps can adapt the function provision and user interface to user needs and enhance the positive user experience.

Health care experts and app designers are often challenged with the question of what functions to offer, as the number of exercise-promotion apps is rapidly increasing [61]. The *attractive* features offered in our Kano study can shed light on the answer, as focusing on attractive quality attributes will outperform only providing expected quality attributes in maintaining strategic advantage [62]. Besides, implementing attractive designs may incur no risk, as low performance in such designs will not increase customer dissatisfaction as defined in the Kano model. One function to build for *all* middle-agers is visual demonstration (A30), as visual demonstration is the sole attractive design to the total sample. Even though the prevalence of this function has been proposed as high as 47% [18], app designers are still encouraged to delve into more varieties of this feature (eg, exercise demonstration with virtual reality [63]). Besides, our analysis also reveals 4 features (ie, A12, A24, A29, and A52) which should be provided to low-PA middle-agers, whereas A11 should be delivered to medium-PA users. This suggestion echoes the importance of behavioral monitoring in PA promotion [34], as A11 (automatically record PA), A12 (set health goals), and A52 (reminding to PA) relates to behavioral facilitation. Furthermore, as motivational interviewing and self-talk were not present in any app analyzed in [33], offering A24 (reminding past success in exercise) would create unique value in the market.

For three design features in demand but with little or no supply (ie, A21, A22, and A23), we propose design guidelines based on existing findings in the literature. For A23 (prompt generalization of exercise), we suggest app designers including indoor maps (eg, Google Indoor Maps) and remind users with location-based messages to encourage stair use. For example, when users are very close to a stair, messages invoking heuristic processing (eg, use the stair) should be used, whereas when a stair is placed at some distance allowing systematic processing, messages should be designed to induce systematic processing such as "will you take the stair? [64]" To perform well in A21 (contingent reward for exercise practice) and A22 (contingent rewards with grading), an app should collect information regarding user practicing PA and offer users with rewards contingent on user behavior. As existing studies have proposed diversified reward structures [65,66], app designers should implement the reward structures and test their effectiveness in a natural setting. Apps can exercise persuasive appeals to induce intrinsic motivation so that users can understand the importance of exercise. This requirement is also reflected in our finding, as A1 (general consequences of exercise) and A2 (customized consequences of exercise) are categorized as *must-be* by low-PA participants. As with extrinsic motivation, app designers can consider to cooperate with advertising agencies and provide users with economic incentives (eg, The AIR MILES incentives [67]).

Although comparing needs against features supplied, social comparison and social support were surprisingly rated as *indifferent*, even in the high-MPSE subsample. This finding seems contrary to the popularity of social networking functions in exercise-promotion apps for the younger population. As most of the participants used social networking apps (eg, LINE and Facebook), we therefore assumed that they should be aware of social networking apps and own experiences in basic functions, which suggests that middle-aged adults did not intend to receive social support or conduct social comparison via exercise-promotion apps. Possible explanations are as follows: for those middle-agers who overlap PA with social life, they may have formed own styles to interact with exercise partners. On the other hand, a habit of separating exercise from social life also weakens the need for sociability of exercise-promotion apps. Accordingly, this study suggests that app developers should consider keeping social networking functionality to a minimum extent, and allowing users to disable functions or hide related widgets in the interface from being seen. Another recommendation is to connect with existing social networking apps (eg, LINE and Facebook) in order to minimize the cognitive load in learning new apps.

Comparison With Prior Work

The formative study was reported by Rabin and Bock [21], who proposed that mobile phone users had a number of specific preferences with regard to PA. This study is in line with [21] in increasing understanding on how mobile phone users perceive app features promoting PA. Uniquely, this study is grounded with the BCTs proposed by Michie et al [22] and targets the

cohort of middle-aged adults, therefore offering more insights on their appraisals on mobile phone features promoting PA. To our knowledge, this work is also the first to address middle-agers' quality perception toward design features of exercise-promotion apps. With our empirical findings, this study offers strategic recommendations for app developers to create value with attractive features that can induce positive emotion.

Existing works have proposed that technology self-efficacy can influence perceived usefulness on computing devices in health care [68]. This study is in concordance with [68] in understanding the effects of technology self-efficacy on perception toward technology use. Uniquely, our finding contributes to current literature in the finding that MPSE influences quality categorization. In particular, high self-efficacy seems to make features look more *attractive*, whereas medium self-efficacy only considers something she or he must have. This finding not only contributes to the self-efficacy studies with new evidence, but also informs practitioners of the importance of increasing user's confidence.

Limitations

The limitations of this study include the small sample size. The findings are also limited in that the participants were not based on probability sampling. As we interpreted the Kano analysis results in an aggressive approach, the effect of *indifferent* perception was therefore weakened through the analysis. A larger sample should be used to alleviate these issues in future study design.

To avoid participant fatigue, we created the questionnaire based on the 40 BCTs proposed by Michie et al [22]. Future studies are recommended to adopt the 93 techniques in the BCT taxonomy (v1) to generate more design features. This study is also limited due to its cross-sectional design. With randomized controlled trials, a study can more easily attribute any observed effect to the treatments being compared, from which strong evidence can be derived [69].

Conclusions and Future Research

Patient-centered care (PCC) advocates that patient needs and preferences should be respected [70]. Following the PCC principle, this study fills the research gap by offering the design recommendations of exercise-promotion mobile phone apps for middle-agers. Visual demonstration is the sole feature that should be implemented for middle-agers, whereas design features customized for middle-aged adults of different characteristics are also provided. By comparing the needs in our findings to the current supply of app features, attractive design features are suggested to enhance strategic advantage of app developers. In addition to these recommended app features, MPSE is identified as a dominant factor inducing attractive and one-dimensional quality perception, where quality categorization by high-MPSE participants mostly (ie, 12/14) falls into *one-dimensional*, whereas all (6/6) of the non- *indifferent* categories by medium MPSE participants are *must-be*. The results by the multinomial regression analysis also indicated a similar pattern. Although current literature have proposed that self-efficacy influences perceived usefulness and perceived ease of use [68], the relationship between technology-use self-efficacy and quality categorization may remain an open question. This calls for future research to explore the underlying mechanisms behind the findings.

Acknowledgments

The authors are grateful for the financial support from Chang Gung Memorial Hospital (CMRPD3D0071 and BMRP745). The authors thank the anonymous reviewers and the editor for their helpful comments on an earlier version of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The scoring protocol of IPAQ-Taiwan. IPAQ: International Physical Activity Questionnaire.

[PDF File (Adobe PDF File), 199KB - mhealth_v5i5e65_app1.pdf]

Multimedia Appendix 2

Kano analysis results.

[PDF File (Adobe PDF File), 762KB - mhealth_v5i5e65_app2.pdf]

Multimedia Appendix 3

Model fitting information and the results of likelihood ratio tests.

[PDF File (Adobe PDF File), 207KB - mhealth_v5i5e65_app3.pdf]

References

1. Ayalon L, Doron I, Bodner E, Inbar N. Macro- and micro-level predictors of age categorization: results from the European Social Survey. Eur J Ageing 2013 May 22;11(1):5-18. [doi: 10.1007/s10433-013-0282-8]
2. UN.org. 2013. World Population Ageing 2013 URL: http://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2013.pdf [WebCite Cache ID 6lZ07dNaw]
3. Driskell J. Nutrition and Exercise Concerns of Middle Age. Boca Raton, Florida: CRC Press; 2009.
4. Lahti J, Holstila A, Mänty M, Lahelma E, Rahkonen O. Changes in leisure time physical activity and subsequent disability retirement: a register-linked cohort study. Int J Behav Nutr Phys Act 2016 Sep 06;13(1):99 [FREE Full text] [doi: 10.1186/s12966-016-0426-2] [Medline: 27599466]
5. Lahti J, Lahelma E, Rahkonen O. Changes in leisure-time physical activity and subsequent sickness absence: a prospective cohort study among middle-aged employees. Prev Med 2012 Dec;55(6):618-622. [doi: 10.1016/j.ypmed.2012.10.006] [Medline: 23064133]
6. Wannamethee SG, Shaper AG, Walker M. Changes in physical activity, mortality, and incidence of coronary heart disease in older men. Lancet 1998 May 30;351(9116):1603-1608. [doi: 10.1016/S0140-6736(97)12355-8] [Medline: 9620713]
7. Godbey G, Chow H, Burnett-Wolle S. New ideas for promoting physical activity among middle age and older adults. J phys educ recreat dance 2007 Sep;78(7):22-26. [doi: 10.1080/07303084.2007.10598053]
8. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
9. Allender S, Cowburn G, Foster C. Understanding participation in sport and physical activity among children and adults: a review of qualitative studies. Health Educ Res 2006 Dec;21(6):826-835. [doi: 10.1093/her/cyl063] [Medline: 16857780]
10. Thorsteinsen K, Vittersø J, Svendsen GB. Increasing physical activity efficiently: an experimental pilot study of a website and mobile phone intervention. Int J Telemed Appl 2014;2014:746232 [FREE Full text] [doi: 10.1155/2014/746232] [Medline: 24963290]
11. Nielsen. The Mobile Consumer: A Global Snapshot URL: http://www.nielsen.com/content/dam/corporate/uk/en/documents/Mobile-Consumer-Report-2013.pdf [WebCite Cache ID 6lF5KYRg2]
12. Bert F, Giacometti M, Gualano MR, Siliquini R. Smartphones and health promotion: a review of the evidence. J Med Syst 2014 Jan;38(1):9995. [doi: 10.1007/s10916-013-9995-7] [Medline: 24346929]
13. Stevens CJ, Bryan AD. Rebranding exercise: there's an app for that. Am J Health Promot 2012;27(2):69-70. [doi: 10.4278/ajhp.120711-CIT-338] [Medline: 23113774]
14. 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;14(2):e55 [FREE Full text] [doi: 10.2196/jmir.1950] [Medline: 22522112]
15. Moxey A, Robertson J, Newby D, Hains I, Williamson M, Pearson S. Computerized clinical decision support for prescribing: provision does not guarantee uptake. J Am Med Inform Assoc 2010;17(1):25-33 [FREE Full text] [doi: 10.1197/jamia.M3170] [Medline: 20064798]
16. Venkatesh V, Thong J, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2013 Mar;36(1):157-178 [FREE Full text]
17. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478 [FREE Full text] [doi: 10.2307/30036540]
18. Koo C, Wati Y, Park K, Lim MK. Website quality, expectation, confirmation, and end user satisfaction: the knowledge-intensive website of the Korean National Cancer Information Center. J Med Internet Res 2011 Nov;13(4):e81 [FREE Full text] [doi: 10.2196/jmir.1574] [Medline: 22047810]
19. Ralston JD, Revere D, Robins LS, Goldberg HI. Patients' experience with a diabetes support programme based on an interactive electronic medical record: qualitative study. Br Med J 2004 May 15;328(7449):1159 [FREE Full text] [doi: 10.1136/bmj.328.7449.1159] [Medline: 15142919]
20. van Gemert-Pijnen JE, 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;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
21. Rabin C, Bock B. Desired features of smartphone applications promoting physical activity. Telemed J E Health 2011 Dec;17(10):801-803. [doi: 10.1089/tmj.2011.0055] [Medline: 22010977]
22. 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]
23. Conroy DE, Yang C, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
24. Witell L, Löfgren M, Dahlgaard JJ. Theory of attractive quality and the Kano methodology--the past, the present, and the future. Total Qual Manag Bus 2013 Dec;24(11-12):1241-1252. [doi: 10.1080/14783363.2013.791117]
25. Shyu J, Chang W, Ko H. Comparative analysis of experience-oriented customer needs and manufacturer supplies based on the Kano model. Total Qual Manag Bus 2013 Dec;24(11-12):1272-1287. [doi: 10.1080/14783363.2012.746196]

26. Chen C, Chuang M. Integrating the Kano model into a robust design approach to enhance customer satisfaction with product design. Int J Prod Econ 2008 Aug;114(2):667-681. [doi: 10.1016/j.ijpe.2008.02.015]
27. Wang C, Wu C. Combining conjoint analysis with Kano model to optimize product varieties of smart phones: a VIKOR perspective. Journal of Industrial and Production Engineering 2014 May 16;31(4):177-186. [doi: 10.1080/21681015.2014.918566]
28. Palumbo F, Dominici G, Basile G. Designing a mobile app for museums according to the drivers of visitor satisfaction. : WSEAS Press; 2013 Jun Presented at: Proceedings of the 1st International Conference on Management, Marketing, Tourism, Retail, Finance and Computer Applications; 2013/6/25-27; Dubrovnik, Croatia p. 159-166.
29. Sulisworo D, Maniquiz N. Integrating Kano's Model and SERVQUAL to improve healthcare service quality. 2012 Presented at: International Conference on Green World in Business and Technology; 2012/3/23-24; Maret, Indonesia p. 130-144.
30. Chang W, Chang Y. Patient satisfaction analysis: identifying key drivers and enhancing service quality of dental care. J Dent Sci 2013 Sep;8(3):239-247. [doi: 10.1016/j.jds.2012.10.006]
31. Dominici G, Palumbo F. How to build an e-learning product: factors for student/customer satisfaction. Bus Horiz 2013 Jan;56(1):87-96. [doi: 10.1016/j.bushor.2012.09.011]
32. Shahin A, Zairi M. Kano model: a dynamic approach for classifying and prioritising requirements of airline travellers with three case studies on international airlines. Total Qual Manag Bus 2009 Sep;20(9):1003-1028. [doi: 10.1080/14783360903181867]
33. 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;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
34. Yang C, Maher JP, Conroy DE. Implementation of behavior change techniques in mobile applications for physical activity. Am J Prev Med 2015 Apr;48(4):452-455. [doi: 10.1016/j.amepre.2014.10.010] [Medline: 25576494]
35. Camerini L, Schulz PJ. Effects of functional interactivity on patients' knowledge, empowerment, and health outcomes: an experimental model-driven evaluation of a web-based intervention. J Med Internet Res 2012;14(4):e105 [FREE Full text] [doi: 10.2196/jmir.1953] [Medline: 22810046]
36. King AC, Rejeski WJ, Buchner DM. Physical activity interventions targeting older adults. A critical review and recommendations. Am J Prev Med 1998 Nov;15(4):316-333. [Medline: 9838975]
37. Lindblom K, Gregory T, Wilson C, Flight IH, Zajac I. The impact of computer self-efficacy, computer anxiety, and perceived usability and acceptability on the efficacy of a decision support tool for colorectal cancer screening. J Am Med Inform Assoc 2012;19(3):407-412 [FREE Full text] [doi: 10.1136/amiajnl-2011-000225] [Medline: 21857024]
38. Compeau DR, Higgins CA. Computer self-efficacy: development of a measure and initial test. MIS Q 1995 Jun;19(2):189 [FREE Full text] [doi: 10.2307/249688]
39. Delbaere K, Hauer K, Lord SR. Evaluation of the incidental and planned activity questionnaire (IPEQ) for older people. Br J Sports Med 2010 Nov;44(14):1029-1034. [doi: 10.1136/bjsm.2009.060350] [Medline: 19474003]
40. Deng HB, Macfarlane DJ, Thomas GN, Lao XQ, Jiang CQ, Cheng KK, et al. Reliability and validity of the IPAQ-Chinese: the Guangzhou Biobank Cohort study. Med Sci Sports Exerc 2008 Feb;40(2):303-307. [doi: 10.1249/mss.0b013e31815b0db5] [Medline: 18202571]
41. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the international physical activity questionnaire short form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011;8:115 [FREE Full text] [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
42. Liou YM, Jwo CJ, Yao KG, Chiang L, Huang L. Selection of appropriate Chinese terms to represent intensity and types of physical activity terms for use in the Taiwan version of IPAQ. J Nurs Res 2008 Dec;16(4):252-263. [Medline: 19061172]
43. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
44. Marshall AL, Leslie ER, Bauman AE, Marcus BH, Owen N. Print versus website physical activity programs: a randomized trial. Am J Prev Med 2003 Aug;25(2):88-94. [Medline: 12880874]
45. Spittaels H, De Bourdeaudhuij I, Vandelanotte C. Evaluation of a website-delivered computer-tailored intervention for increasing physical activity in the general population. Prev Med 2007 Mar;44(3):209-217. [doi: 10.1016/j.ypmed.2006.11.010] [Medline: 17197015]
46. Naugle KM, Riley JL. Self-reported physical activity predicts pain inhibitory and facilitatory function. Med Sci Sports Exerc 2014 Mar;46(3):622-629 [FREE Full text] [doi: 10.1249/MSS.0b013e3182a69cf1] [Medline: 23899890]
47. Park JE, Lee J, Kim B, Kim KW, Chae SH, Cho MJ. Above-moderate physical activity reduces both incident and persistent late-life depression in rural Koreans. Int J Geriatr Psychiatry 2015 Jul;30(7):766-775. [doi: 10.1002/gps.4244] [Medline: 25503946]
48. Heesch KC, van Uffelen JG, Hill RL, Brown WJ. What do IPAQ questions mean to older adults? Lessons from cognitive interviews. Int J Behav Nutr Phys Act 2010 May 11;7:35 [FREE Full text] [doi: 10.1186/1479-5868-7-35] [Medline: 20459758]

49. Mikulić J, Prebežac D. A critical review of techniques for classifying quality attributes in the Kano model. Managing Service Quality 2011 Jan 25;21(1):46-66. [doi: 10.1108/09604521111100243]
50. Yang C. The refined Kano's model and its application. Total Qual Manag Bus 2007 Jan 24;16(10):1127-1137. [doi: 10.1080/14783360500235850]
51. Xu Q, Jiao RJ, Yang X, Helander M, Khalid HM, Opperud A. An analytical Kano model for customer need analysis. Design Studies 2009 Jan;30(1):87-110. [doi: 10.1016/j.destud.2008.07.001]
52. Svedberg P, Lichtenstein P, Pedersen NL. Age and sex differences in genetic and environmental factors for self-rated health: a twin study. J Gerontol B Psychol Sci Soc Sci 2001 May;56(3):S171-S178. [Medline: 11316842]
53. Ohayon MM, Carskadon MA, Guilleminault C, Vitiello MV. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep 2004 Nov 01;27(7):1255-1273. [Medline: 15586779]
54. Sinharay S, Stern HS, Russell D. The use of multiple imputation for the analysis of missing data. Psychol Methods 2001;6(4):317-329. [doi: 10.1037/1082-989X.6.4.317]
55. Chen L. A novel approach to regression analysis for the classification of quality attributes in the Kano model: an empirical test in the food and beverage industry. Omega 2012 Oct;40(5):651-659. [doi: 10.1016/j.omega.2011.12.004]
56. Lee M, Newcomb J. Applying the Kano methodology to meet customer requirements: NASA's microgravity science program. Quality Management Journal 1997 Apr;4(3):95-106 [FREE Full text]
57. Townsend N, Wickramasinghe K, Williams J, Bhatnagar P, Rayner M. Physical Activity Statistics 2015. London, UK: British Heart Foundation; Jan 28, 2015.
58. Sun F, Norman IJ, While AE. Physical activity in older people: a systematic review. BMC Public Health 2013;13:449 [FREE Full text] [doi: 10.1186/1471-2458-13-449] [Medline: 23648225]
59. Hundert AS, Huguet A, McGrath PJ, Stinson JN, Wheaton M. Commercially available mobile phone headache diary apps: a systematic review. JMIR Mhealth Uhealth 2014;2(3):e36 [FREE Full text] [doi: 10.2196/mhealth.3452] [Medline: 25138438]
60. Morita PP, Cafazzo JA. Challenges and paradoxes of human factors in health technology design. JMIR Hum Factors 2016;3(1):e11 [FREE Full text] [doi: 10.2196/humanfactors.4653] [Medline: 27025862]
61. Khalaf S. Flurrymobile.tumblr. Health and fitness apps finally take off, fueled by fitness fanatics URL: http://flurrymobile.tumblr.com/post/115192181465/health-and-fitness-apps-finally-take-off-fueled [accessed 2016-11-07] [WebCite Cache ID 6lqJuByMC]
62. Chen Y, Su C. A Kano-CKM model for customer knowledge discovery. Total Qual Manag Bus 2006 Jun;17(5):589-608. [doi: 10.1080/14783360600588158]
63. Merchant Z, Goetz ET, Cifuentes L, Keeney-Kennicutt W, Davis TJ. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. Comput Educ 2014 Jan;70:29-40. [doi: 10.1016/j.compedu.2013.07.033]
64. Suri G, Sheppes G, Leslie S, Gross JJ. Stairs or escalator? Using theories of persuasion and motivation to facilitate healthy decision making. J Exp Psychol Appl 2014 Dec;20(4):295-302. [doi: 10.1037/xap0000026] [Medline: 25180946]
65. Hurley JC, Hollingshead KE, Todd M, Jarrett CL, Tucker WJ, Angadi SS, et al. The walking interventions through texting (WalkIT) trial: rationale, design, and protocol for a factorial randomized controlled trial of adaptive interventions for overweight and obese, inactive adults. JMIR Res Protoc 2015 Sep 11;4(3):e108 [FREE Full text] [doi: 10.2196/resprot.4856] [Medline: 26362511]
66. van Mierlo T, Hyatt D, Ching AT, Fournier R, Dembo RS. Behavioral economics, wearable devices, and cooperative games: results from a population-based intervention to increase physical activity. JMIR Serious Games 2016;4(1):e1 [FREE Full text] [doi: 10.2196/games.5358] [Medline: 26821955]
67. Goyal S, Morita PP, Picton P, Seto E, Zbib A, Cafazzo JA. Uptake of a consumer-focused mHealth application for the assessment and prevention of heart disease: the <30 days study. JMIR Mhealth Uhealth 2016;4(1):e32 [FREE Full text] [doi: 10.2196/mhealth.4730] [Medline: 27012937]
68. Wu J, Wang S, Lin L. Mobile computing acceptance factors in the healthcare industry: a structural equation model. Int J Med Inform 2007 Jan;76(1):66-77. [doi: 10.1016/j.ijmedinf.2006.06.006] [Medline: 16901749]
69. Barton S. Which clinical studies provide the best evidence? The best RCT still trumps the best observational study. Br Med J 2000 Jul 29;321(7256):255-256 [FREE Full text] [Medline: 10915111]
70. Lambert BL, Street RL, Cegala DJ, Smith DH, Kurtz S, Schofield T. Provider-patient communication, patient-centered care, and the mangle of practice. Health Commun 1997 Jan;9(1):27-43. [doi: 10.1207/s15327027hc0901_3]

Abbreviations

BCT: behavior change technique
BMI: body mass index
CALO-RE: Coventry, Aberdeen, and London—Refined
CGMH: Chang Gung Memorial Hospital

CPU: central processing unit
IPAQ: International Physical Activity Questionnaire
IPAQ-LF: International Physical Activity Questionnaire-long form
IPAQ-SF: International Physical Activity Questionnaire-short form
IRB: Institutional Review Board
IT: information technology
OR: odds ratio
PA: physical activity
PCC: patient-centered care
SMS: short message service
UTAUT: unified theory of acceptance and use of technology

Edited by G Eysenbach; submitted 23.06.16; peer-reviewed by M Khazaeepool, M Fiordelli, P Lee; comments to author 14.09.16; revised version received 11.11.16; accepted 24.11.16; published 25.05.17.

Please cite as:
Liao GY, Chien YT, Chen YJ, Hsiung HF, Chen HJ, Hsieh MH, Wu WJ
What to Build for Middle-Agers to Come? Attractive and Necessary Functions of Exercise-Promotion Mobile Phone Apps: A Cross-Sectional Study
JMIR Mhealth Uhealth 2017;5(5):e65
URL: http://mhealth.jmir.org/2017/5/e65/
doi:10.2196/mhealth.6233
PMID:28546140

Original Paper

An mHealth App for Supporting Quitters to Manage Cigarette Cravings With Short Bouts of Physical Activity: A Randomized Pilot Feasibility and Acceptability Study

Mary Hassandra[1], MSc, PhD; Taru Lintunen[1], Lic Psych, MSc, PhD; Martin S Hagger[1,2,3], PhD; Risto Heikkinen[4], MS (Statistics); Mauno Vanhala[5,6], MD; Tarja Kettunen[6,7], MSc, PhD

[1]Faculty of Sport and Health Sciences, University of Jyvaskyla, Jyvaskyla, Finland

[2]School of Psychology and Speech Pathology, Faculty of Health Sciences, Health Psychology & Behavioral Medicine Research Group & Laboratory of Self-Regulation (Laser), Curtin University, Perth, Australia

[3]Department of Physical Education, Hong Kong Baptist University, Hong Kong, China (Hong Kong)

[4]Statistical Analysis Services, Analyysitoimisto Statisti Oy, Jyvaskyla, Finland

[5]School of Medicine, Unit of Primary Health Care, University of Eastern Finland, Kuopio, Finland

[6]Unit of Primary Health Care, Central Finland Health Care District, Jyvaskyla, Finland

[7]Faculty of Sport and Health Sciences, Health Sciences, University of Jyvaskyla, Jyvaskyla, Finland

Corresponding Author:
Mary Hassandra, MSc, PhD
Faculty of Sport and Health Sciences
University of Jyvaskyla
PO Box 35 University of Jyvaskyla
Jyvaskyla, 40014
Finland
Phone: 358 0408053979
Fax: 358 0408053979
Email: maria.m.chasandra@jyu.fi

Abstract

Background: While gains in reducing smoking rates in Finland have been made, prevalence rates are still substantial. Relapse rates among smokers engaged in quit-smoking programs are high. Physical activity has been proposed as one means to help smokers manage cravings. Software and apps on mobile phone and handheld devices offer an opportunity to communicate messages on how to use physical activity to manage cravings as part of quit-smoking programs.

Objective: We aimed to test the feasibility, acceptability, usability, and preliminary efficacy of an mHealth mobile phone app, Physical activity over Smoking (PhoS), to assist smokers in quitting smoking in a randomized controlled trial. The app was designed to prompt smokers to engage in physical activities to manage their smoking cravings.

Methods: Regular smokers (n=44) attended a group-based behavioral counselling program aimed at promoting physical activity as an additional aid to quit. After quit day, participants were randomly allocated to an intervention (n=25) or to a comparison (n=19) group. Participants in the intervention group were provided with the PhoS app and training on how to use it to assist with relapse prevention. Participants in the comparison condition were provided with generalized relapse prevention training.

Results: Some participants reported that the PhoS app was useful in assisting them to successfully manage their cigarette cravings, although compliance across the sample was modest and participants reported low levels of usability. Participants receiving the PhoS app did not report greater abstinence than those who did not receive the app. However, participants receiving the app were more likely to report greater abstinence if they did not use pharmacological support, while those who did not receive the app reported greater abstinence when using pharmacological support. Participants receiving the app reported greater levels of physical activity than those who did not. Results revealed that the app resulted in better retention.

Conclusions: The PhoS app showed some potential to reduce abstinence among participants not using pharmacological therapy and to increase physical activity. However, problems with usability and lack of effects on abstinence raise questions over the app's long-term effectiveness. Future research should prioritize further development of the app to maximize usability and test effects of the intervention independent of quit-smoking programs.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 55259451; http://www.controlled-trials.com/ISRCTN55259451 (Archived by WebCite at http://www.webcitation.org/6cKF2mzEI)

(JMIR Mhealth Uhealth 2017;5(5):e74) doi:10.2196/mhealth.6252

KEYWORDS

behavior change; mHealth app; physical activity; randomized controlled trial; relapse prevention; smoking

Introduction

The harmful effects of smoking on health are well documented, but quit rates are low and long-term relapse rates range from 75%, for smokers adopting combined therapies of counseling and pharmacotherapy to assist quitting, to 95%, for smokers who adopt a complete abstinence ("cold turkey") strategy without pharmacological or therapeutic support [1-3]. A systematic review of the effectiveness of smoking relapse prevention interventions [4] indicated that self-help treatments can aid relapse prevention. Mobile phone apps have been used either as a stand-alone treatment program to assist quitting or as a self-help tool in combination with other treatment programs [5]. Behavioral interventions delivered through mobile phones and handheld devices offer promising opportunities to expand psychological practice [6,7], and the integration of effective behavior change interventions using software in these devices [8] (collectively known as mHealth apps) may help develop stronger evidence in future research in this field.

Physical activity has recently been incorporated into existing smoking cessation counselling programs as a cessation aid [9-12]. There is good evidence for the acute, short-term effects of physical activity on smoking-related variables. Research has indicated that physical activity acutely reduces cigarette craving [13-15]. The literature suggests a wide range of intensities, from isometric exercises and yoga to high-intensity activities with heart rates of up to 85% of maximum, as an aid to managing cigarette cravings and withdrawal symptoms [13-17]. However, although evidence of the positive effects of physical activity on reducing cravings is promising [13-16], most of the research focuses on acute, short-term effects, and findings are limited by a lack of long-term follow-up and a focus on laboratory-based studies. Therefore, there is a need for interventions to examine the long-term effects of physical activity on cigarette cravings and smoking cessation in real-life situations. Similar efforts in this line of research have been initiated [17,18], but none have adopted mobile technology and mHealth apps.

The use of mHealth apps may provide a valuable platform to test the feasibility of physical activity interventions to reduce smoking cravings and promote quitting in real-life contexts. To serve this purpose, we developed an mHealth app, Physical activity over Smoking (PhoS), to support smokers trying to quit in managing their cigarette cravings and abstaining in the long term using short bouts of physical activity as a behavioral substitution strategy. We based the PhoS app on psychological theory and evidence-based relapse prevention strategies [19]. Following its development, we aimed to assess the feasibility and acceptability of the PhoS app for use by quitting smokers, and to test its effects in changing outcome measures (abstinence, self-reported cravings, relapses, and awareness of cravings) in a small-scale randomized controlled trial (ISRCTN55259451; Multimedia Appendix 1 [20]).

The aim of this preliminary study was to assess the feasibility, acceptability, and preliminary efficacy of the newly developed mHealth app PhoS. We predicted that use of the app would be feasible as a means to assist smokers wanting to quit and would have high levels of acceptability consistent with research using mHealth apps in other behavioral health contexts. We also hypothesized that smokers using the app would have high abstinence rates, lower self-reported cravings and relapses, and greater awareness of cravings.

Methods

Participants, Procedure, and Randomization

We screened participants for eligibility by having them complete a short battery of questions online. Eligibility criteria were age (adults 18-65 years), regular smoking (smoking >10 cigarettes/day), no existing mental health condition or other addictions, expression of high motivation to quit, and no health conditions for which physical activity was contraindicated. Noneligible participants were referred to their general practitioner for further health advice. We invited participants who met the eligibility criteria and gave verbal consent to participate, via a telephone call, to take part in a smoking cessation program, consisting of 3 weekly counseling sessions, and helped them set a quit date. Participants who reached the quit day were randomly allocated to an intervention or a comparison group. Within 3 to 7 days after their quit day, participants in both groups had a fourth session with the study nurse and were provided with group training on relapse prevention and prompted to form action plans to cope with their cravings. Participants in the intervention group received additional training for using the mHealth app. After they downloaded the PhoS app to their mobile phones, they were provided with instructions on how to use it as an additional support tool whenever they experienced cravings. Participants without a mobile phone or with one that was incompatible with the app were provided with a mobile phone with the PhoS app preloaded for the duration of the study. Those allocated to the comparison group were provided with guidance to develop an action plan to overcome relapses. The intervention group received the same training but with an emphasis on identifying short bouts of physical activity that can be used in everyday-life situations to overcome cravings. Then, they were provided with instruction on the use of the PhoS app as a support tool.

We randomly allocated participants to the intervention or the comparison group using version 4 of Research Randomizer, an online randomization tool [21]. The principal investigator generated the allocation sequence and the study nurse assigned

participants to groups. The principal investigator was blinded to the participants' data up to the first follow-up occasion, and the study nurse was blinded to the allocation sequence until the day of randomization. Participants were blinded to group assignment. Sessions took place at the Jyvaskyla Community Primary Health Care Center in Central Finland and the University of Jyvaskyla campus.

Screening data were collected via an online questionnaire. Those who were eligible received a telephone call from the study nurse to set up a time for the first session at the health center or university campus to complete baseline measures after they agreed to participate by signing the consent form. We collected data from the face-to-face sessions via a paper-and-pencil questionnaire. Follow-up data were collected via online questionnaires after an invitation from the research assistant via email or text message, which contained a link to the online questionnaire. App users' data were automatically uploaded daily to a secure university server reserved for this purpose. We obtained ethics approval from the Ethics Committee of the Central Finland Health Care District (Keski-Suomen Sairaanhoitopiirin Eettinen Toimikunta).

Intervention

The intervention has been described in detail in the published study protocol [19]. The main theoretical frameworks used for the intervention and app development were sociocognitive theory [22], the theory of planned behavior [23], control theory [24], the relapse prevention model [25], and motivational interviewing [26]. The behavior change techniques based on these theoretical approaches that we used in the interventions were self-monitoring, setting goals, setting graded tasks, reviewing behavioral goals, and planning action and coping.

The majority of the sessions were group based, but we organized some individual sessions for those who missed the group sessions. All sessions were delivered by the study nurse, who had previous experience with counseling smokers to quit and an additional 3 hours of training from the research group on how to deliver the specific intervention. In most of the sessions, a member of the research group who was also an accredited psychologist was also present and kept notes throughout the procedure to ensure adherence to the protocol and participated in group discussions related to the promotion of physical activity if needed.

The first 3 sessions, common to both groups, aimed to help participants quit smoking. The promotion of physical activity as a supportive behavior to decrease, and eventually quit, smoking was a central part of these sessions. The aim of the fourth session, which was held after random allocation to the intervention and comparison groups, was to support participants in developing relapse prevention plans and action plans to manage cravings. The only difference between the 2 groups in this fourth session was that, for intervention group there, we emphasized using short bouts of physical activity (taking a brisk walk, stretching, breathing, balance, etc) as the main craving management strategy.

The PhoS app was introduced as a tool that would give them ideas of what physical activities to do and how to do them depending on their individual and situational conditions. The app drew from a database that contained strategies based on the relapse prevention model [27,28] and the taxonomy of behavior change techniques for smoking cessation [29]. The app database includes a pool of 57 introductory messages (eg, "Researchers have found 70 poisonous chemicals in cigarettes which cause cancer. Stay healthy!"), 49 motivational messages (eg, "feel like a winner, not miserable after 3 minutes!"), and 64 physical activities (eg, "Stretch your upper arms. Hold for 10 seconds each"), all of which were coded to appear according to the users' profile and status.

Data entries were self-initiated whenever participants used the app. All usage data were automatically uploaded to the server whenever the device was connected to the Internet. Participants' identification number and profile settings were entered when the PhoS app was installed on their mobile phones and used for the first time. After that, every time users reported experiencing a craving, they were prompted to provide situational information regarding mood, place, and social environment. According to users' profile information and the situational status, a variety of physical activities accompanied by theory-based motivational messages were suggested. Physical activities were animated to demonstrate the suggested activities in a visual form. Finally, each time participants used the app, they were asked to provide feedback if they had successfully managed the craving. Screenshot 1 in Multimedia Appendix 2 provides details of the profile settings (eg, sex, age, days since quit date) screen. Screenshot 2 provides an example of the situational data users could enter each time they used the app (mood, place, and social environment). Screenshot 3 displays some examples of the animated physical activities provided to participants to demonstrate the activities. Finally, screenshot 4 provides an example of the feedback users were given after they had used the app.

Participants' use of the app was monitored by the PhoS app, which automatically uploaded participants' data to a central server. When participants had not used the PhoS app for 1 week, they were contacted to check whether there was a technical problem or they had stopped using the app. If they reported that they were not using the app any longer, they answered a short questionnaire to assess the reasons for discontinuing its use.

Measures

Implementation Outcomes

We assessed engagement with the PhoS app through data extracted from participants' mobile phones indicating frequency and duration of app use. Moreover, we calculated retention rates of the app users for the same time points as the other follow-up measures.

We assessed fidelity of app use through open- and closed-ended questions to record whether participants had used the app as instructed, as well as questions regarding any other additional support they used (eg, pharmacological, other mHealth apps). Additional usage monitoring data were extracted from the phones to describe usage. Fidelity was assessed at 3 days and at 1, 3, and 6 months after participants had started using the app.

We assessed usability of the app twice, at 1 week and 1 month, after participants had started using the app. Participants completed the System Usability Scale (SUS) self-reported questionnaire [30], a 10-item scale measuring subjective components of usability. Responses were provided on 5-point scales ranging from strongly agree (1 point) to strongly disagree (5 points). Moreover, participants who stopped using the app at any point during the trial were prompted to supply reasons why they had done so using an open-ended question.

Preliminary Efficacy Outcomes

The primary efficacy outcome was self-reported 7-day point prevalence abstinence (PPA) at 7 days prior to each scheduled follow-up [31]. PPA measures after the quit time point and at the end of the follow-up (6 months) were verified by a saliva cotinine test [32].

Secondary outcome measures were self-reported number of relapses during the last 7 days, self-reported number of cravings during the last 7 days, and self-efficacy of being aware of experiencing cravings (AEF; "How well are you aware of your cigarette cravings?"). These 3 single-question assessments were collected at the same time points as the main outcome measure. We assessed 2 more secondary outcomes: self-efficacy in managing cravings (MCEF; "How well do you manage your cravings?") and the power of control in managing cravings (CCM). The CCM was assessed with 6 items with different permutations for the intervention and control groups. An example item is "If I am in a situation where I celebrate with my friends...," then the answer for the intervention group was "It will be more difficult to use [the] PhoS app to control my craving for tobacco," and for the comparison group it was "It will be more difficult to do something to control my craving for tobacco." Responses were given on 7-point scales ranging from totally agree (1 point) to totally disagree (7 points). Multimedia Appendix 3 displays the intervention timeline and the detailed measurement time points for each group and all measures.

We also included additional outcome measures: self-reported physical activity assessed by the International Physical Activity Questionnaire (IPAQ) [33] collected at sessions 1 and 3, and at the 6 month follow-up time points; and attitude, intention, and perceived behavioral control with respect to increasing physical activity behavior [34] at sessions 1 and 4. We assessed attitude using 6 items on 7-point semantic differential scales ranging from –3 to +3 points (eg, "To increase physical activity in the following month will be for me..." good-bad, unpleasant-pleasant, wise-silly, easy-difficult, healthy-unhealthy, important-not important). We measured intention using 3 items with responses provided on 7-point scales ranging from very likely (1 point) to very unlikely (7 points) (eg, "In the following month I intend to increase physical activity"). We measured perceived behavioral control using 4 items on 7-point scales ranging from complete control (1 point) to very little control (7 points) (eg, "How much control will you exert over exercising regularly during the next month?").

Statistical Analyses

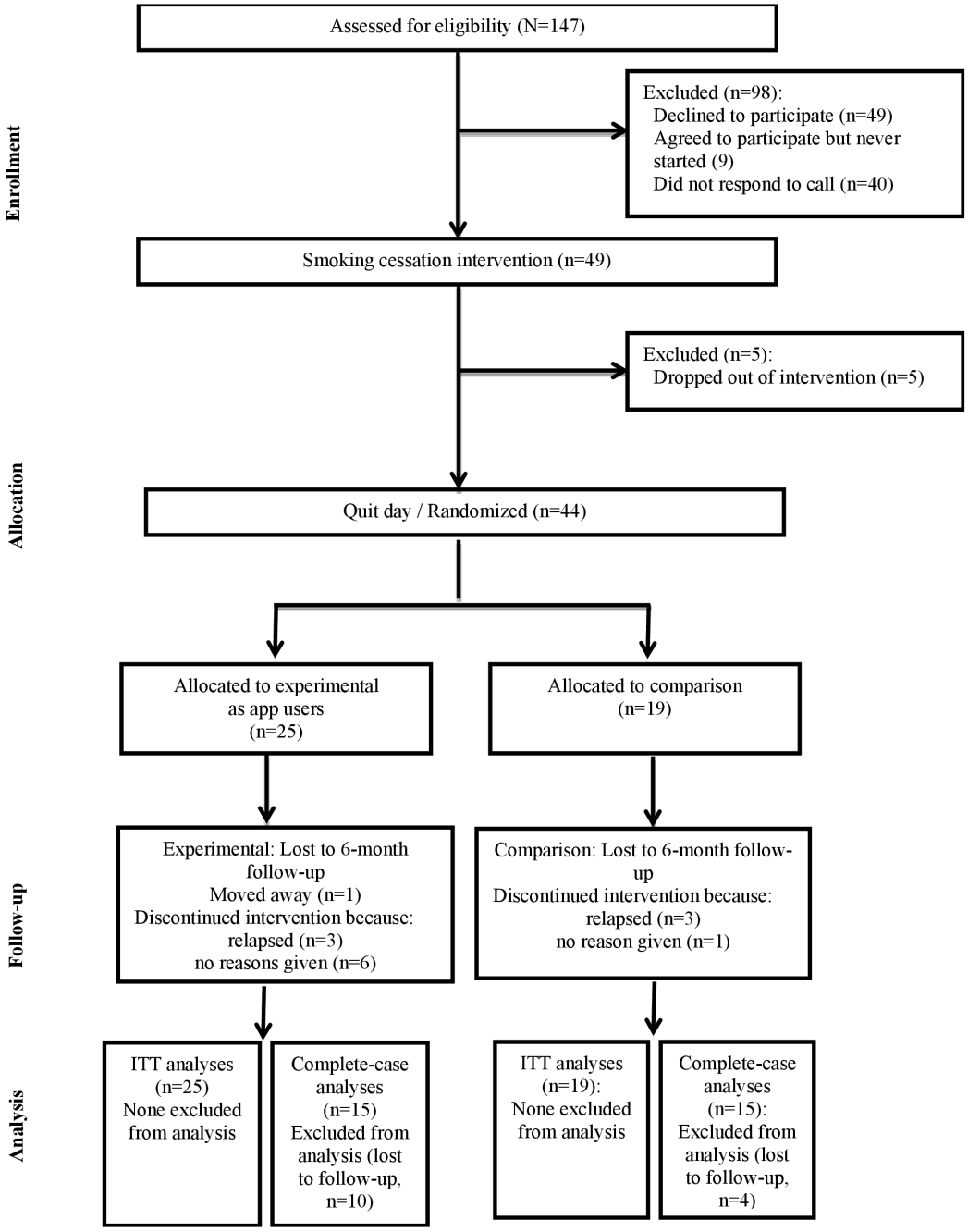
We conducted all main analyses using the intention-to-treat method with participants remaining in their originally assigned groups after random allocation regardless of adherence or protocol deviation. We also performed the same analyses using complete-case analysis. We report data using descriptive statistics, including mean, standard deviation, and frequencies, and used analysis of variance to test group comparisons on baseline and demographic data. Logistic regression examined differences between the groups on the dichotomous primary outcome variable: abstinence versus nonabstinence at the 6-month follow-up time point. We tested group differences at all follow-up time points using a 3 (time points: baseline, 1 week, 6 months) × 2 (groups: intervention, control) generalized linear mixed model (GLMM) with logit link function. We identified the use of pharmacological support as an important control variable in the preliminary analyses and added it to the models as a covariate. We also used a series of GLMMs of identical design to analyze the continuous secondary efficacy measures AEF, MCEF, and CCM during the follow-up period. In addition, a series of GLMMs with Poisson link function using the same design as the previous analyses analyzed the number of cigarettes and number of cravings. For number of relapses, AEF, MCEF, and CCM, nonparametric Mann-Whitney *U* tests compared differences between the intervention and comparison groups at the 6-month time point. We used R statistical package software version 3.1.3 (R Foundation) for all analyses.

Results

Participant Flow

The initial response rates in recruitment were lower than expected, so we extended the recruitment period and used several additional recruitment methods. A radio interview with the study nurse on a local radio station was the most successful and immediate recruitment strategy. Of the 147 individuals who completed the screening assessment and were eligible, 49 agreed to participate and entered the smoking cessation intervention. A total of 44 participants reported quitting and were randomly allocated into intervention (n=25) and comparison (n=19) groups. The groups were not balanced because the group allocation was an ongoing procedure and the randomization sequence was generated for 50 participants (25 for each group), but only 44 people were finally randomly allocated. Figure 1 illustrates the flow of participants. Some participants did not attend all 3 of the prequit sessions. However, we included those who attended at least one session, managed to quit, and also attended the fourth session in the follow-up group (n=4). A total of 34 participants completed all of the follow-up measures (77% retention rate); 19 of the 25 participants allocated to the intervention group completed all the follow-up measures (76% retention rate) and 15 of the 19 participants allocated to the comparison group completed all the follow-up measures (79% retention rate). The most common reason participants gave for dropping out at follow-up was relapse back to smoking and no longer wanting to continue the study.

Figure 1. Participant flow chart. ITT: intention-to-treat.



Baseline Data

All participants were of Finnish nationality. Six (14%) participants had a university education, 16 (36%) had an applied sciences education, 14 (31%) had a vocational school education, and 3 (7%) had a secondary school education only. Regarding employment status, 33 (75%) were employed, 3 (7%) were unemployed, 3 (7%) were pensioners, and 5 (11%) were university students. Table 1 presents additional characteristics of the study participants.

Table 2 presents descriptive measures by group for all sessions. There were no differences between groups in demographic characteristics (Table 1) or among all measures during all sessions (Table 2).

Table 1. Participant characteristics[a].

Characteristics	Study group					
	All (n=44)		Intervention (n=25)		Comparison (n=19)	
	Mean or n	SD or %	Mean or n	SD or %	Mean or n	SD or %
Age in years, mean (SD)	39.11	10.67	39.92	11.16	38.05	11.16
Salary (€), mean (SD)	28,613	11,436.30	29,280	10,121.22	27,736	13,207.07
BMI[b]in kg/m^2, mean (SD)	25.53	3.89	25.35	3.91	25.78	3.97
Pharmacological support						
Used	29	66	16	64	13	68
Not used	15	34	9	36	6	32
Sex						
Male	25	57	15	60	10	53
Female	19	43	10	40	9	47

[a]No significant between-group differences were detected in the means.

[b]BMI: body mass index.

Table 2. Descriptive measures (mean scores) during all sessions by group.

Measures	Study group								
	All			Intervention			Comparison		
	n	Mean	SD	n	Mean	SD	n	Mean	SD
Session 1 (baseline)									
Smoking behavior 7 days[a]	41	15.35	7.30	22	14.87	8.69	19	15.90	5.44
IPAQ[b]	41	3853.24	4719.79	22	4481.18	5527.46	19	3126.15	3580.26
Attitude physical activity	41	1.69	.45	22	1.68	.45	19	1.71	.46
Intention physical activity	41	1.79	.69	22	1.72	.66	19	1.86	.74
Perceived behavioral control over physical activity	41	1.95	.64	22	1.98	.70	19	1.90	.58
CCM[c]	41	3.23	1.24	22	3.28	1.42	19	3.17	1.04
Session 2									
Smoking behavior 7 days	38	11.08	4.41	19	11.42	5.07	19	10.75	3.73
Session 3									
Smoking behavior 7 days	37	8.96	4.19	19	9.35	4.27	18	8.55	4.18
IPAQ	36	3981.66	4865.02	19	3204.10	3770.73	17	4850.70	5853.08
CCM	37	3.60	1.34	19	3.39	1.48	18	3.82	1.17
Session 4									
Smoking behavior 7 days	32	8.45	4.19	17	8.18	4.16	15	8.77	4.35
Attitude physical activity	44	1.49	.66	25	1.54	.78	19	1.43	.47
Intention physical activity	37	2.36	1.32	21	2.65	1.55	16	2.00	.85
Perceived behavioral control over physical activity	37	2.25	.97	21	2.41	1.20	16	2.04	.51

[a]Data from the everyday self-monitoring diary.

[b]IPAQ: International Physical Activity Questionnaire.

[c]CCM: power of control in managing cravings.

Of the 44 participants, 38 took the saliva cotinine test, and 11 of them were verified as abstinent (index 0 or 1), based on a cutoff value of 10 ng/mL, which is the stated abstinence using NicAlert saliva cotinine tests [32]. A total of 19 of them had a medium cotinine index (2-4), and 8 had an index score of 5 or 6. All participants with a score higher than 1 stated that they were using cotinine replacements (2- to 4-mg tablets or gum 2-8 times/day) and therefore were considered as eligible to enter the follow-up period for measurements. Data on the saliva cotinine test at the 6-month follow-up were not available. We offered the participants several options to take the test (home visit, by mail, etc), but uptake was extremely low (n=5). Therefore, verification of self-reported data at the end was not possible.

Implementation Outcomes

Engagement

Data from some mobile phones could not be uploaded to the server when connected to the Web, which left 14 (56%) participants' data available for engagement and compliance analysis. We contacted participants to offer them a new phone, but they declined. We omitted phone data from 5 participants from analysis because they did not use the app at all after the initial session. We selected only fully completed cases of use of the app from the start button entry to the feedback button entry. We excluded incomplete data units, as we considered them not to be valid uses of the app. Complete phone data were extracted for 9 users. Table 3 displays the average use data at the same time points as the other follow-up measures.

Table 3. Average frequencies of use during the period measured at each time point.

Use	Time point							
	3 days	1 week	2 weeks	3 weeks	4 weeks	8 weeks	12 weeks	24 weeks
Duration of use (days)[a]	1.56	0.56	1.11	0.78	1.00	3.56	2.22	2.56
No. of days of use[b]	1.44	0.56	0.67	0.56	0.33	0.78	0.56	0.44
Average uses/day[c]	1.09	0.78	0.44	0.44	0.22	0.22	0.42	0.30
Total uses[d]	2.00	0.78	0.67	0.56	0.33	0.78	1.33	0.67
No. of minutes of use[e]	1.79	0.69	0.58	0.80	0.12	0.70	0.97	0.76
Views[f]	12.22	3.89	4.11	9.67	3.67	11.56	8.67	10.67

[a]Time period between the first and last day of using the app for each period of measurement.

[b]Number of different days the participant used the app.

[c]Number of times on average the user used the app in a single day.

[d](Number of days) × (average number of uses).

[e]Number of minutes the participant used the app in total

[f]Number of different views the user has had on the screen during use.

Figure 2. Proportion of app users during follow-up.

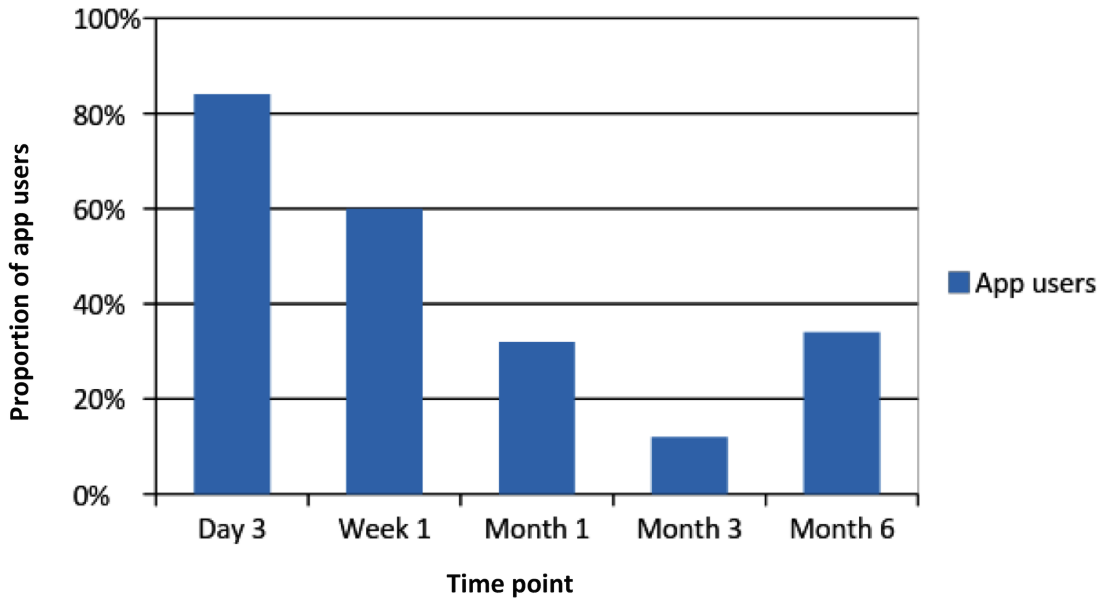


Figure 2 displays the retention rates for app use throughout the follow-up period. The retention rate at day 3 was the highest and at month 3 it was the lowest. We defined retention rate as the ratio of the number of retained users of the app to the number of participants who completed all study measures at the same time point.

Fidelity

Fidelity questions for the intervention group aimed to identify whether participants used the app, how and when they used it, and, if they decided not use it anymore, the reasons why. Figure 3 illustrates their responses to the closed-ended questions.

Participants' (n=5) reasons for using the app were "just for fun" or "just browsing," while reasons they stopped using it were "After the initial excitement I don't think it is for me" or "I don't need it any more."

Fidelity questions for the comparison group tried to identify whether participants used any kind of additional support to manage their cravings during the follow-up period. Figure 4 illustrates their responses to the closed-ended questions.

Participants' responses to the open-ended question asking about what other methods they used when they faced cravings were as follows: pharmacological support (eg, Champix, nicotine replacement products); snus (smokeless tobacco); exercise, jogging, or running; walks in nature; and doing housework.

Table 4 displays phone data regarding the average frequencies of participants' status (mood, place, social environment) and relapse reports at every measurement time point. Figure 5 illustrates the overall average frequency of situation status and relapse reporting of all app users, based on phone data, during the 6-month follow-up.

Usability

We assessed perceived quality of the PhoS app twice, at weeks 1 and 4, after participants had started using the app. Participants' (n=23) average SUS score 1 week after they started using the app was 21.6 out of 100. Participants' (n=20) scores were 21.8 at 1-month follow-up. This indicates a relatively low level of usability, considering that SUS scores below 68 are assumed to be below average [30]. Participants' reasons for stopping using the app were that (1) they were trying to avoid using their phone outside of work time, (2) the app did not work on their phone and carrying a second phone for the express purpose of using the app was inconvenient, (3) some perceived the suggested tasks to be "weird," (4) they did not feel the need to use the app and they could manage cravings without it, (5) it was not powerful enough to help them overcome their cravings, and (6) they did not use apps in general.

Table 4. Average frequencies of situation status and relapse reporting of app users at every measurement time point during follow-up.

Status	Time point							
	3 days	1 week	2 weeks	3 weeks	4 weeks	8 weeks	12 weeks	24 weeks
Positive feedback[a]	0.89	0.44	0.22	0.11	0.00	0.22	1.11	0.56
Neutral feedback[b]	1.11	0.33	0.44	0.44	0.33	0.56	0.22	0.11
Negative feedback[c]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Outdoors	0.22	0.00	0.00	0.11	0.00	0.00	0.00	0.00
At work	0.44	0.11	0.00	0.00	0.11	0.00	0.00	0.00
At home	1.33	0.67	0.67	0.44	0.22	0.78	1.33	0.67
Alone	0.89	0.44	0.22	0.44	0.11	0.22	0.00	0.44
Not alone	1.11	0.33	0.44	0.11	0.22	0.56	1.33	0.22
Positive mood	0.67	0.22	0.11	0.00	0.00	0.22	0.33	0.56
Neutral mood	1.00	0.44	0.56	0.44	0.33	0.44	0.78	0.11
Negative mood	0.33	0.11	0.00	0.11	0.00	0.11	0.22	0.00

[a]"The app was helpful to manage craving."

[b]"I managed not to relapse, but the app didn't help."

[c]"I relapsed."

Figure 3. Fidelity responses of the intervention group detailing when, how, and why participants used the app.

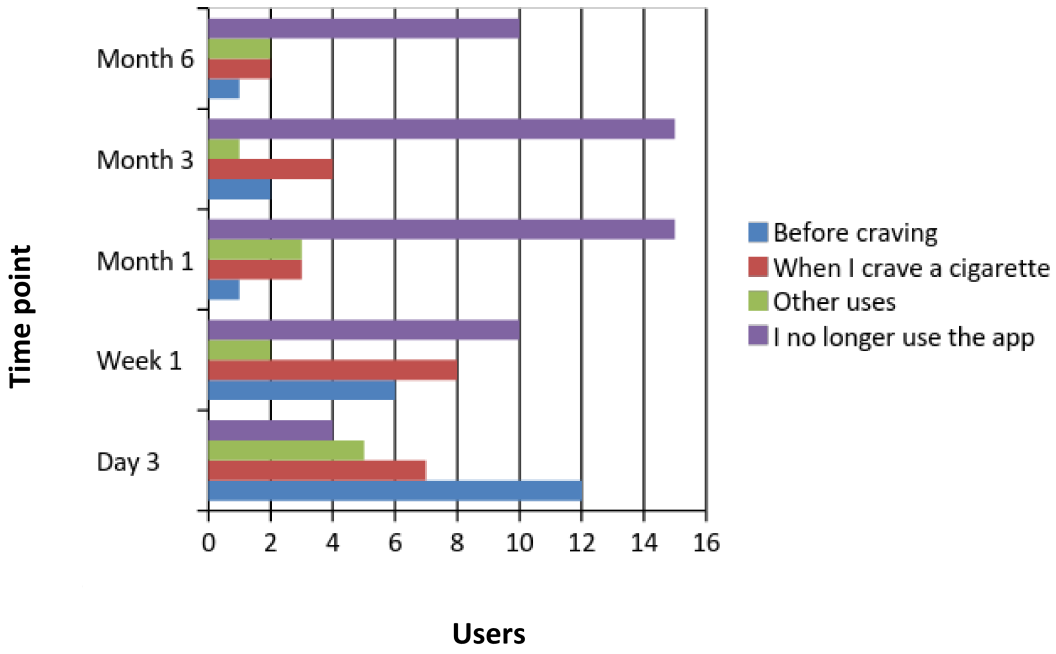


Figure 4. Fidelity responses of the comparison group detailing whether they used additional support.

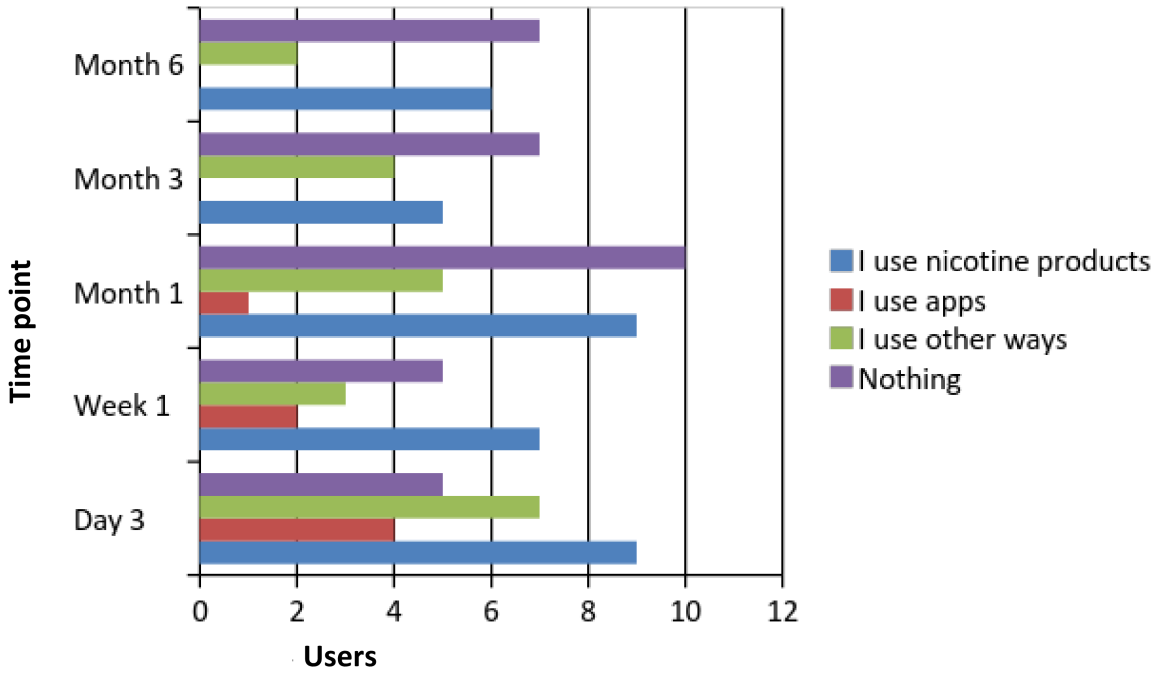
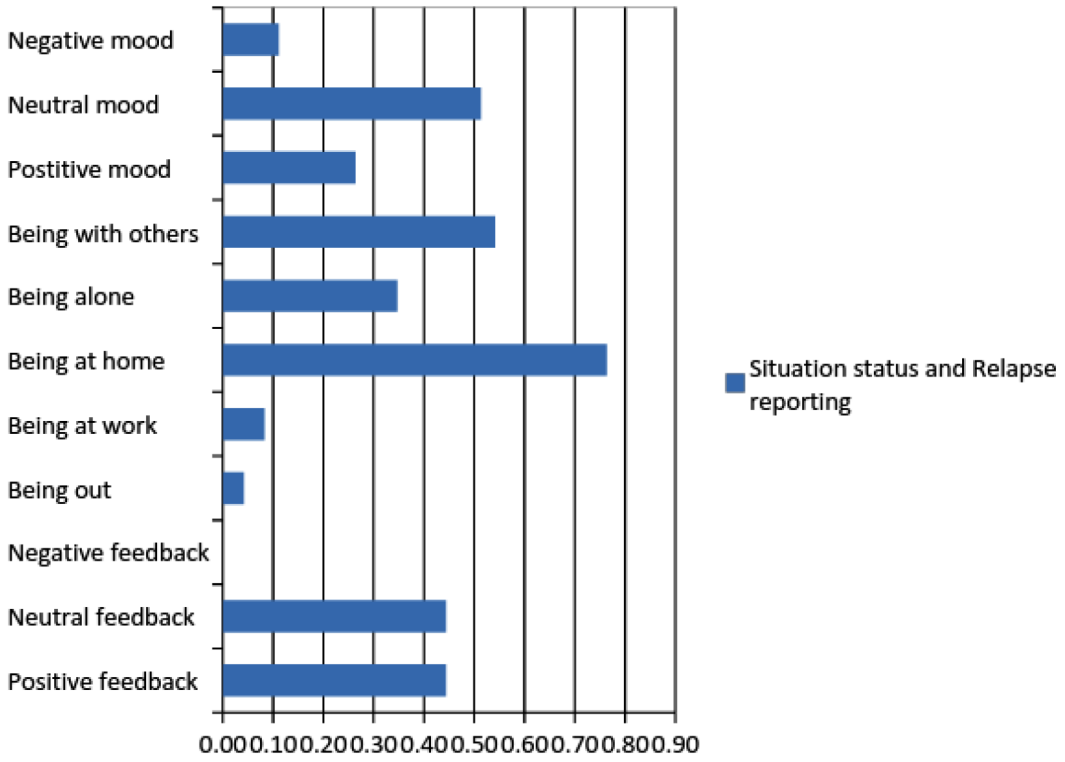


Figure 5. Average frequency of situation status and relapse reporting of all app users during the 6-month follow-up, based on phone data.



Preliminary Efficacy Outcomes

Overall, intention-to-treat analyses revealed that 36% (n=16) of the 44 participants who entered the follow-up period as quitters remained abstinent after 6 months. The abstinence proportion increased to 53% for those who provided complete data at all time measurement points (n=30) up to 6 months.

A logistic regression analysis examined the effects of intervention group (intervention vs control) for differences in the 7-day PPA at the 6-month follow-up period. Odds of quitting were not statistically significantly different between the groups for the intention-to-treat (n=44) and complete-case analyses (n=30). Multimedia Appendix 4 displays the frequencies and percentages of successful quitters during the 6-month follow-up period by group for the intention-to-treat and complete-case analyses.

We performed the same logistic regression analysis including use of pharmacological support at the end of the follow-up period as a covariate, since pharmacological support is strongly related to the main outcome. Table 5 presents descriptive measures. Intention-to-treat analysis revealed that the odds of quitting in the comparison group were lower for those who did not use any pharmacological support (odds ratio [OR] 0.23, 95% CI 0.02-2.59, *P*=.24), while odds of quitting in the intervention group were higher for those who did not use pharmacological support (OR 16.07, 95% CI 0.83-313). The group × pharmacological support interaction effect fell just short of statistical significance (χ^2_{40}=53.7, *P*=.05).

Table 5. Frequencies and percentages of successful quitters at the 6-month follow-up time point by group and use of pharmacological support (nonabstinent/abstinent) for intention-to-treat and complete-case analyses.

Study group	Intention-to-treat		Complete case	
	Nonabstinent	Abstinent	Nonabstinent	Abstinent
Intervention group, n (%)				
Nonabstinent	12 (75)	4 (44)	5 (56)	1 (17)
Abstinent	4 (25)	5 (56)	4 (44)	5 (83%)
Total	16 (64)	9 (36)	9 (60)	6 (40)
Total (nonabstinent + abstinent)	25 (100)		15 (100)	
Comparison group, n (%)				
Nonabstinent	7 (54)	5 (83)	3 (33)	5 (83)
Abstinent	6 (46)	1 (17)	6 (67)	1 (17)
Total	13 (68)	6 (32)	9 (60)	6 (40)
Total (nonabstinent + abstinent)	19 (100)		15 (100)	

The interaction effect was statistically significant in the complete-case analysis (χ^2_{26}=34.6, *P*=.01), indicating that, in the intervention group, the odds of quitting were higher for those who did not use pharmacological support (OR 62.8, 95% CI 1.73-2259, *P*=.02). In the comparison group, the odds of quitting were lower for those who did not use any pharmacological support (OR 0.10, 95% CI 0.0-11.29, *P*=.08). A 2 (intervention: intervention vs comparison) × 7 (time points: follow-ups at 1, 2, 3, 4, 8, 12, and 24 weeks postintervention) GLMM examined the effects of PhoS app use versus nonapp use for differences in 7-day PPA during the 6-month follow-up period. There was no statistically significant interaction between time trend and intervention group (χ^2_1=3.6, *P*=.06) for the intention-to-treat group. There was a statistically significant interaction effect between intervention group and use of pharmacological support (χ^2_1=4.1, *P*=.04), indicating that those participants in the intervention group who did not use pharmacological support were more likely to be abstinent during the entire follow-up period.

Secondary Outcomes

Mann-Whitney *U* tests indicated that there were no statistically significant differences between groups for all secondary outcome measures (relapses, cravings, AEF, MCEF, and CCM) at the end of the follow-up period (6 months). Similarly, GLMMs indicated that there were no statistically significant differences between groups for the same secondary outcome measures during the entire follow-up period. Multimedia Appendix 5 displays the descriptive statistics of all secondary measures at all follow-up time points by group.

We included all secondary measures as covariates in the analysis examining the effect of the intervention and time point on the primary efficacy measure, abstinence proportion. Only AEF at the end of the 6-month follow-up period and the intervention group had a significant interaction effect (χ^2_1=4.2, *P*=.04) on abstinence for the complete-case analysis. More specifically, in the intervention group AEF increased abstinence levels, meaning that for every unit increase in AEF the odds of quitting were multiplied by 2.5. However, in the comparison group AEF decreased abstinence, meaning that for every unit increase in AEF the odds of quitting were multiplied by 0.77.

Additional Outcomes

Physical Activity Behavior

Mann-Whitney *U* test for differences in IPAQ scores between session 3 and at the end of the 6-month follow-up period revealed significant differences (*U*=23, *P*=.008) between the intervention and comparison groups. This result indicates that the distribution of the difference in IPAQ scores between those time points was significantly different across groups. Intervention group participants reported higher physical activity scores (mean rank=16.20, n=10) than comparison group members (mean rank=8.77, n=13). The same analysis for testing the distribution of the difference in IPAQ scores between session 1 and the 6-month follow-up period revealed no statistically significant differences across groups (*U*=77, *P*=.55).

Theory of Planned Behavior Constructs of Physical Activity Behavior

We tested the role of the sociocognitive variables from the theory of planned behavior for physical activity in predicting secondary outcomes (number of cigarettes smoked, AEF, CCM, and MCEF) using linear multiple regression. Specifically, the dependent variable of interest was regressed on the theory of planned behavior variables controlling for treatment group, pharmacological support use, and time by using intention-to-treat and complete-case analyses. We conducted separate analyses for each dependent variable for complete cases. The number of cigarettes smoked was a dependent variable at all follow-up time points. Perceived behavioral control over physical activity was a statistically significant predictor of number of cigarettes smoked (B=–5.27, SE 2.18, *P*=.01) and MCEF (B=1.89, SE 0.57, *P*=.003). The negative coefficient indicated that a higher perceived behavioral control over physical activity score was related to fewer cigarettes smoked during all follow-up periods. An identical analysis using the theory of planned behavior measures taken at session 1

revealed no statistically significant results. There were no other statistically significant effects. The linear mixed-effects model, for complete cases, of MCEF × time was significant, indicating that as time was progressing, MCEF was decreasing (B=–0.75, SE 0.30, *P*=.02).

Discussion

Implementation Outcomes

Use of the PhoS app was lower than expected. Retention rates for app use declined throughout the follow-up period. The increase of retention rate at 6 months (30%) after the lower rate at 3 months (10%) implies that there was a group of stable users throughout the period. After locating those 4 participants, we determined that their probability of quitting compared with the comparison group was higher. Characteristics common to these participants were their annual income, having a higher education degree, not using any pharmacological support during the study, and considerably increasing their physical activity from baseline to the end of the 6-month follow-up period.

Intervention group participants reported that they used the app mostly when they were at home, with others, and in a neutral mood. Relapse reporting indicated that when they used the app they were able to successfully manage their cravings and not relapse. Nevertheless, the frequency of positive and neutral feedback indicates that the app was helpful for some participants but not others for relapse prevention. None of the phone entries reported a relapse. Qualitative data from the fidelity check questions revealed when, how, and why participants used, or did not use, the app. Most participants who reported using the app used it to plan what to do either before experiencing a craving or when they had actually experienced a craving. However, some reported that they used the app for other reasons also (for fun or just browsing).

Usability

Usability level results were very low. The low SUS scores indicate that there is a need to review the app and identify usability problems. However, SUS itself is not diagnostic, so the results do not shed light on the reasons for low reported usability. One of the reasons was that, although participants were given the option to replace their phones with study mobile phones if the app could not operate on their own device, for several reasons some participants did not accept that. Using a second phone for accessing the app introduced an additional potential bias of usability. The open question for usability revealed some reasons for the low scores. Most of them are common reasons that apply to most apps: the person did not use apps in general or there were compatibility issues with their device. Fidelity questions regarding the reasons for not using the app were informative for the attrition in app use. For example, 1 participant reported that “After the initial excitement I don’t think it’s for me,” suggesting that the novelty decreases, and sustained engagement with these kinds of apps is low and has been previously reported as a reason for attrition in the use of mHealth apps. Research has indicated that users’ initial interest in mHealth apps quickly fades as the novelty wears off [35]. Nevertheless, statements such as “I don’t need it anymore” were made by users who stopped smoking, used the app for some time (as long as they experienced cravings), and then stopped using it once their cravings faded because there was no reason to continue. Finally, a few statements indicated that some suggested physical activities were “weird” or were not powerful enough to help participants overcome their cravings. Overall, there is significant room for improvement in increasing the novelty and usability of the PhoS app to improve sustained engagement and use.

Preliminary Efficacy Outcomes

We found that intervention participants provided with the PhoS app and training on how to use it did not report greater abstinence or differences in outcomes relating to craving management relative to the comparison group. The lack of differences between the groups suggests that the app did not provide added value in assisting quitting or managing cravings. This result may be attributable to the low frequency of use and level of usability of the app. A further possible explanation may be that we tested the PhoS app against a strong comparison group in which participants received the same counseling smoking cessation program that was promoting physical activity as a means to manage cigarette cravings. The lack of any additional added effect of the app might be attributed to the relative strength of the effect of the counseling intervention in the comparison group. Responses to fidelity questions by comparison group participants implied that they also used several forms of physical activity as a method to manage their cravings (eg, exercise, jogging, or running, walks in nature, or housework). It would be interesting in a future study of the app to use a usual-care comparison group to test whether the app alone can have an added effect on abstinence rates.

The overall abstinence for all participants (n=44), after 6 months, was 36% in intention-to-treat analyses and 53% in complete-case analyses. These long-term abstinence rates are considered satisfactory, but while generalized comparisons can be made with other smoking cessation programs, no direct comparison can be made given that this intervention included several unique components: behavioral counselling to quit smoking, physical activity promotion, and pharmacological support. A previous study [36] reported abstinence rates of 42% (for cognitive behavior therapy plus nicotine replacement therapy) and 36% (for physical activity plus nicotine replacement therapy) at 12-month follow-up. Overall, our results compare favorably with these abstinence rates and support the use of physical activity as a means to manage smoking cravings in smoking cessation programs [13-15].

Pharmacological support was also an important moderator of the effects of the intervention on abstinence. Odds of quitting were lower among participants in the comparison group who did not use any pharmacological support, while odds of quitting were higher in the intervention group for those who did not use pharmacological support. These interactive effects were evident during the entire follow-up period. This finding implies that the app may be an effective supportive tool to promote abstinence instead of using pharmacological support. Although participants in the intervention group were free to choose to use pharmacological support if they wanted to, they were not asked why they did not use pharmacological support. Therefore, it is

not clear whether this finding is caused by chance, since it was not possible to locate a similar finding in previous research, so this finding should be treated as preliminary. An interesting avenue for future research would be to examine whether mHealth apps can be used as an alternative means of support quitting smoking for people who cannot use, or do not want to use, pharmacological support.

Secondary Outcomes

After testing all 5 secondary measures as covariates to the primary efficacy measure, there was a significant interaction between AEF and group, indicating that increased AEF and being a member of the intervention group increased the odds to stay abstinent. According to self-efficacy theory [37,38], developing awareness of specific situations where efficacy may be low and mentally rehearsing desired behavior in these situations appears to enhance efficacy for behavior change. This finding paves the way for future interventions promoting quitting smoking to identify situations where participants might lapse and visualize potential solutions or responses to direct them away from the typical response of lighting up a cigarette.

Additional Outcomes

Although there were no differences between the groups of users and nonusers of the app in the main outcome, there were important group differences in some of the additional outcome measures. The differences between groups on self-reported physical activity behavior suggest that the app acted as a reminder of being more physically active as well. The question arising from this finding is whether any physical activity promotion app would have the same effect. According to a meta-analysis [39], mobile devices are an effective means for influencing physical activity behavior, but no study has tested the effect of these apps on smoking cessation.

In addition, higher perceived behavioral control over physical activity behavior, measured at session 4, was related to participants smoking fewer cigarettes and higher MCEF capacity at all follow-up periods, regardless of group allocation and pharmacological support. Since the same result did not apply to perceived behavioral control over physical activity behavior at session 1, it is likely that such changes were attributable to the smoking cessation intervention. Sociocognitive theories postulate that sense of control is the most powerful source of self-efficacy [22,40]. The possible mechanism behind this might be that successful attempts to control physical activity behavior had the effect of strengthening self-efficacy to manage cigarette cravings. Theorists support the existence of this mechanism, but they claim that this is possible mainly for the same behavior or domain [22,41], whereas, based on our data, this mechanism worked for another behavior. This is consistent with research demonstrating that sociocognitive beliefs translate across health behavior contexts [42]. Nevertheless, the focus of this study and the low number of participants precluded a test for a causal relationship and should be a priority for future research.

Strengths and Limitations

To our knowledge, this is the first study to develop and test the feasibility, acceptability, usability, and preliminary efficacy of an mHealth mobile phone app that promotes physical activity as a means to manage smoking cravings. Lessons learned from this study may inform future research in this area. The small number of participants was an a priori limitation of this pilot study identified in the study protocol [19]. In addition, as is common to much research on smoking cessation and mHealth apps, retention was one of the main challenges faced in this study [35,43,44], and this had a direct effect on the power of the study results [45,46]. Moreover, generalizability and applicability are limited due to the low use and usability of the app.

Future Research

We developed the PhoS app for research purposes, so the focus was on developing an app with a strong basis in theory- and evidence-based behavior change techniques. An equal focus on design, attractiveness, and usability would increase use. Therefore, future attempts to use the PhoS app as a tool should first focus on improving its usability. For that reason, the app code is available, upon request from the first author, for research purposes only.

Introducing the PhoS app earlier to the participants, during the smoking cessation intervention, when participants decrease the number of cigarettes and increase their physical activity, might improve use of the app (eg, more consistent and coherent use, at the right time and for the right reasons). This way there would be more time to practice using the app and provide opportunities to receive feedback on its use from those delivering the intervention.

Testing the PhoS app among people who have quit smoking with a method or program that does not use physical activity as a supportive aid might reveal useful information in a future trial. For example, an important question is whether the app can stand alone and help quitters manage their cravings through short bouts of physical activity when it is not used in conjunction with another intervention, such as the counselling quit-smoking intervention in this study. Another direction for future research should be to test the use of a physical activity promotion app as a supportive relapse prevention tool in quit-smoking interventions.

Conclusion

Overall, implementation results from this study indicated that the PhoS app needs improvements before being embedded in a larger trial. Moreover, the app did not have an additive effect on abstinence based on smoking abstinence beyond the effects of the counselling quit-smoking intervention. Nevertheless, findings on the secondary dependent variables shed some light on the possible mechanisms that were activated by its use, independently and in combination with the counselling and pharmacological smoking cessation intervention. Despite the challenges of research on mHealth apps, there is documented potential and research on developing apps that are optimally engaging and usable, as well as an evidence from behavioral science, that these apps may assist in this potential being realized in the domain of smoking cessation.

Acknowledgments

The study was funded by the Finnish National Institute for Health & Welfare (Terveyden ja Hyvinvoinnin Laitos). We gratefully acknowledge support from the Jyvaskyla Community Primary Health Care Center).

Authors' Contributions

MH, MV, TL, and TK conceived the original idea for the trial and also sought and obtained funding. The manuscript was written by MH with input from all coauthors. MH contributed to the additional analyses and the revision of the paper. RH analyzed the data. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.1.

[PDF File (Adobe PDF File), 547KB - mhealth_v5i5e74_app1.pdf]

Multimedia Appendix 2

Physical activity over Smoking (PhoS) app screenshots.

[PNG File, 330KB - mhealth_v5i5e74_app2.png]

Multimedia Appendix 3

Measures timeline.

[PDF File (Adobe PDF File), 58KB - mhealth_v5i5e74_app3.pdf]

Multimedia Appendix 4

Successful quitters.

[PDF File (Adobe PDF File), 16KB - mhealth_v5i5e74_app4.pdf]

Multimedia Appendix 5

Secondary outcomes descriptive statistics.

[PDF File (Adobe PDF File), 16KB - mhealth_v5i5e74_app5.pdf]

References

1. Hughes JR, Keely J, Naud S. Shape of the relapse curve and long-term abstinence among untreated smokers. Addiction 2004 Jan;99(1):29-38 [FREE Full text] [Medline: 14678060]
2. Fiore M, Jaen C, Baker T. Treating Tobacco Use and Dependence: 2008 Update. Rockville, MD: U.S. Department of Health and Human Services. Public Health Service; 2008.
3. Ferguson J, Bauld L, Chesterman J, Judge K. The English smoking treatment services: one-year outcomes. Addiction 2005 Apr;100(Suppl 2):59-69. [doi: 10.1111/j.1360-0443.2005.01028.x] [Medline: 15755262]
4. Agboola S, McNeill A, Coleman T, Leonardi BJ. A systematic review of the effectiveness of smoking relapse prevention interventions for abstinent smokers. Addiction 2010 Aug;105(8):1362-1380. [doi: 10.1111/j.1360-0443.2010.02996.x] [Medline: 20653619]
5. Aguilera A, Muench F. There's an app for that: information technology applications for cognitive behavioral practitioners. Behav Ther (N Y N Y) 2012 Apr;35(4):65-73 [FREE Full text] [Medline: 25530659]
6. Schueller SM, Munoz RF, Mohr DC. Realizing the potential of behavioral intervention technologies. Curr Dir Psychol Sci 2013 Dec 03;22(6):478-483. [doi: 10.1177/0963721413495872]
7. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]

8. 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]
9. Marcus BH, Albrecht AE, Niaura RS, Abrams DB, Thompson PD. Usefulness of physical exercise for maintaining smoking cessation in women. Am J Cardiol 1991 Aug 1;68(4):406-407. [Medline: 1858687]
10. Marcus BH, Albrecht AE, King TK, Parisi AF, Pinto BM, Roberts M, et al. The efficacy of exercise as an aid for smoking cessation in women: a randomized controlled trial. Arch Intern Med 1999 Jun 14;159(11):1229-1234. [Medline: 10371231]
11. Martin JE, Calfas KJ, Patten CA, Polarek M, Hofstetter CR, Noto J, et al. Prospective evaluation of three smoking interventions in 205 recovering alcoholics: one-year results of Project SCRAP-Tobacco. J Consult Clin Psychol 1997 Feb;65(1):190-194. [Medline: 9103749]
12. Ussher MH, Taylor AH, Faulkner GEJ. Exercise interventions for smoking cessation. Cochrane Database Syst Rev 2014 Aug 29(8):CD002295. [doi: 10.1002/14651858.CD002295.pub5] [Medline: 25170798]
13. Taylor AH, Ussher MH, Faulkner G. The acute effects of exercise on cigarette cravings, withdrawal symptoms, affect and smoking behaviour: a systematic review. Addiction 2007 Apr;102(4):534-543. [doi: 10.1111/j.1360-0443.2006.01739.x] [Medline: 17286639]
14. Roberts V, Maddison R, Simpson C, Bullen C, Prapavessis H. The acute effects of exercise on cigarette cravings, withdrawal symptoms, affect, and smoking behaviour: systematic review update and meta-analysis. Psychopharmacology (Berl) 2012 Jul;222(1):1-15. [doi: 10.1007/s00213-012-2731-z] [Medline: 22585034]
15. Haasova M, Warren FC, Ussher M, Janse Van Rensburg K, Faulkner G, Cropley M, et al. The acute effects of physical activity on cigarette cravings: exploration of potential moderators, mediators and physical activity attributes using individual participant data (IPD) meta-analyses. Psychopharmacology (Berl) 2014 Apr;231(7):1267-1275. [doi: 10.1007/s00213-014-3450-4] [Medline: 24522330]
16. Glass TW, Maher CG. Physical activity reduces cigarette cravings. Br J Sports Med 2014 Aug;48(16):1263-1264. [doi: 10.1136/bjsports-2013-092525] [Medline: 23709520]
17. Fong AJ, De JS, Bray SR, Prapavessis H. Effect of exercise on cigarette cravings and ad libitum smoking following concurrent stressors. Addict Behav 2014 Oct;39(10):1516-1521. [doi: 10.1016/j.addbeh.2014.05.027] [Medline: 24971700]
18. Williams DM, Ussher M, Dunsiger S, Miranda R, Gwaltney CJ, Monti PM, et al. Overcoming limitations in previous research on exercise as a smoking cessation treatment: rationale and design of the “Quit for Health” trial. Contemp Clin Trials 2014 Jan;37(1):33-42 [FREE Full text] [doi: 10.1016/j.cct.2013.11.005] [Medline: 24246818]
19. Hassandra M, Lintunen T, Kettunen T, Vanhala M, Toivonen H, Kinnunen K, et al. Effectiveness of a mobile phone app for adults that uses physical activity as a tool to manage cigarette craving after smoking cessation: a study protocol for a randomized controlled trial. JMIR Res Protoc 2015;4(4):e125 [FREE Full text] [doi: 10.2196/resprot.4600] [Medline: 26494256]
20. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
21. Urbaniak G, Plous S. Research Randomizer. Middletown, CT: Social Psychology Network, Wesleyan University; 2015. URL: https://www.randomizer.org [accessed 2015-10-14] [WebCite Cache ID 6cGlD5rTI]
22. Bandura A. Self-Efficacy: The Exercise of Control. New York, NY: WH Freeman; 1997.
23. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
24. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. Psychol Bull 1982 Jul;92(1):111-135. [Medline: 7134324]
25. Marlatt G, Gordon J. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors. New York, NY: Guilford Press; 1985.
26. Miller W, Rollnick S. Motivational interviewing: preparing people to change addictive behavior. In: Motivational Interviewing: Preparing People to Change Addictive Behavior. New York, NY: Guilford Press; 1991.
27. Hendershot CS, Witkiewitz K, George WH, Marlatt GA. Relapse prevention for addictive behaviors. Subst Abuse Treat Prev Policy 2011;6:17 [FREE Full text] [doi: 10.1186/1747-597X-6-17] [Medline: 21771314]
28. Larimer ME, Palmer RS, Marlatt GA. Relapse prevention. An overview of Marlatt's cognitive-behavioral model. Alcohol Res Health 1999;23(2):151-160. [Medline: 10890810]
29. Michie S, Hyder N, Walia A, West R. Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. Addict Behav 2011 Apr;36(4):315-319. [doi: 10.1016/j.addbeh.2010.11.016] [Medline: 21215528]
30. Brooke J. SUS: a quick and dirty usability scale. 1996. URL: http://cui.unige.ch/isi/icle-wiki/_media/ipm:test-suschapt [accessed 2017-05-17] [WebCite Cache ID 6qWZXWNSu]
31. Hughes JR, Carpenter MJ, Naud S. Do point prevalence and prolonged abstinence measures produce similar results in smoking cessation studies? A systematic review. Nicotine Tob Res 2010 Jul;12(7):756-762 [FREE Full text] [doi: 10.1093/ntr/ntq078] [Medline: 20504946]

32. Cooke F, Bullen C, Whittaker R, McRobbie H, Chen M, Walker N. Diagnostic accuracy of NicAlert cotinine test strips in saliva for verifying smoking status. Nicotine Tob Res 2008 Apr;10(4):607-612. [doi: 10.1080/14622200801978680] [Medline: 18418783]
33. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
34. Francis J, Eccles M, Johnston M, Walker A, Grimshaw J, Foy R, et al. Constructing Questionnaires Based on the Theory of Planned Behaviour: A Manual for Health Services Research. Newcastle upon Tyne, UK: Centre for Health Services Research; 2004.
35. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
36. Prapavessis H, Cameron L, Baldi JC, Robinson S, Borrie K, Harper T, et al. The effects of exercise and nicotine replacement therapy on smoking rates in women. Addict Behav 2007 Jul;32(7):1416-1432. [doi: 10.1016/j.addbeh.2006.10.005] [Medline: 17097814]
37. Bandura A. Self-efficacy mechanism in human agency. Am Psychol 1982;37(2):122-147. [doi: 10.1037/0003-066X.37.2.122]
38. Strecher VJ, DeVellis BM, Becker MH, Rosenstock IM. The role of self-efficacy in achieving health behavior change. Health Educ Q 1986;13(1):73-92. [Medline: 3957687]
39. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res 2012;14(6):e161 [FREE Full text] [doi: 10.2196/jmir.2171] [Medline: 23171838]
40. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
41. Maddux J. Self-efficacy: the power of believing you can. In: Snyder CR, Lopez SJ, editors. Handbook of Positive Psychology. New York, NY: Oxford University Press; 2002:277-287.
42. Hagger MS, Chatzisarantis NLD. The trans-contextual model of autonomous motivation in education: conceptual and empirical issues and meta-analysis. Rev Educ Res 2016 Jun;86(2):360-407 [FREE Full text] [doi: 10.3102/0034654315585005] [Medline: 27274585]
43. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
44. Morgan AJ, Jorm AF, Mackinnon AJ. Internet-based recruitment to a depression prevention intervention: lessons from the Mood Memos study. J Med Internet Res 2013;15(2):e31 [FREE Full text] [doi: 10.2196/jmir.2262] [Medline: 23403043]
45. Fewtrell MS, Kennedy K, Singhal A, Martin RM, Ness A, Hadders-Algra M, et al. How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? Arch Dis Child 2008 Jun;93(6):458-461. [doi: 10.1136/adc.2007.127316] [Medline: 18495909]
46. Sully BG, Julious SA, Nicholl J. A reinvestigation of recruitment to randomised, controlled, multicenter trials: a review of trials funded by two UK funding agencies. Trials 2013;14:166 [FREE Full text] [doi: 10.1186/1745-6215-14-166] [Medline: 23758961]

Abbreviations

AEF: self-efficacy of being aware of experiencing cravings
CCM: power of control in managing cravings
GLMM: generalized linear mixed model
IPAQ: International Physical Activity Questionnaire
MCEF: self-efficacy in managing cravings
OR: odds ratio
PhoS: Physical activity over Smoking
PPA: point prevalence abstinence
SUS: System Usability Scale

Edited by G Eysenbach; submitted 21.06.16; peer-reviewed by J McClure, K Coa, S Hales; comments to author 18.11.16; revised version received 31.03.17; accepted 17.04.17; published 26.05.17.

Please cite as:
Hassandra M, Lintunen T, Hagger MS, Heikkinen R, Vanhala M, Kettunen T
An mHealth App for Supporting Quitters to Manage Cigarette Cravings With Short Bouts of Physical Activity: A Randomized Pilot Feasibility and Acceptability Study
JMIR Mhealth Uhealth 2017;5(5):e74
URL: http://mhealth.jmir.org/2017/5/e74/
doi:10.2196/mhealth.6252
PMID:28550004

Original Paper

Perceptions of Patient Engagement Applications During Pregnancy: A Qualitative Assessment of the Patient's Perspective

Maren Goetz[1]; Mitho Müller[2], Dipl -Psych; Lina Maria Matthies[1], MD; Jenny Hansen[1], B A; Anne Doster[1], MD; Akos Szabo[1], MD; Jan Pauluschke-Fröhlich[3], MD; Harald Abele[3], MD; Christof Sohn[1], PhD; Markus Wallwiener[1*], PhD; Stephanie Wallwiener[1*], MD

[1]Department of Obstetrics and Gynecology, University of Heidelberg, Heidelberg, Germany

[2]Department of Psychology, Ludwig Maximilian University, Munich, Germany

[3]Department of Obstetrics and Gynecology, University of Tübingen, Tübingen, Germany

[*]these authors contributed equally

Corresponding Author:
Stephanie Wallwiener, MD
Department of Obstetrics and Gynecology
University of Heidelberg
Im Neuenheimer Feld 440
Heidelberg, 69121
Germany
Phone: 49 6221 5637551
Fax: 49 6221 5633997
Email: Stephanie.wallwiener@googlemail.com

Abstract

Background: With growing demand for medical information and health applications in pregnancy, the potential of electronic health (eHealth) and mobile health (mHealth) solutions in clinical care is increasingly unfolding. However, we still do not know how pregnant women engage with mobile apps, how such apps impact routine medical care, and whether benefit expectations are met. Whereas recent research has raised the subject of user distribution and analyzed the content of pregnancy applications, there is still a significant knowledge gap regarding what pregnant women like and dislike about pregnancy tools, along with how such interventions could be improved.

Objective: The aim of the study was to examine the perceptions and expectations of mobile and Web-based patient-engagement pregnancy applications. We assessed usability requirements, general acceptance of eHealth, and the impact of eHealth and mHealth pregnancy applications on the doctor-patient interaction and daily clinical routine.

Methods: A qualitative study was conducted at the maternity department of a major German university hospital. The sample included 30 women with low- to medium-risk pregnancies. Half of the patients were seen during outpatient care and half were hospitalized for several days. The extent and frequency of Web- and mobile phone app usage were assessed. Semistructured interviews were conducted and analyzed using systematic thematic analysis.

Results: Patients had a high demand for Web-based pregnancy applications. Study findings suggested a strong request for personalization, monitoring, and accessibility for frequent use as main themes derived from the interviews. Fostering patient empowerment in the doctor-patient relationship was also highly valued for a pregnancy app. Participants favored further integration of medical apps in their daily routine and pregnancy care. However, concerns were raised about content quality, trustworthiness of Web sources, and individual data security.

Conclusions: eHealth and mHealth applications are a highly frequented source of information. Expectations and usability requirements for those applications are also high, thus posing a challenge to interdisciplinary service providers. Patients' attitude toward integrating apps in routine care settings was positive with a favorable influence on patient empowerment. Health care professionals should guide pregnant women toward a successful integration of these educational tools in pregnancy care.

(JMIR Mhealth Uhealth 2017;5(5):e73) doi:10.2196/mhealth.7040

KEYWORDS

pregnancy; telemedicine; mobile applications; information procurement; patient participation; qualitative research

Introduction

With patients' growing demand for medical information and, at the same time, the rapidly evolving opportunities for health care providers to integrate and adopt information technology, the potential of electronic health (eHealth) and mobile health (mHealth) solutions is increasingly unfolding [1]. Here, eHealth is seen as the interface of medical informatics, health care research, and health services, offered preferably via the Web or mobile technology [2,3]. mHealth encompasses the use of mobile communication and multimedia, and their integration in wireless health care delivery systems [4,5]. There is growing evidence that such applications and interventions enhance the doctor-patient connection toward a more partner-like relationship, leading to "patient empowerment" and "patient engagement" [6-11]. Providing evidence-based information and Web-based access to electronic health reports could open new pathways for informing patients. In addition to the omnipresence and penetration of mobile phones in our society, these options have the potential to avail mHealth for prenatal and newborn care, especially in developing economies [12].

Women of childbearing age frequently use the Web and mobile phone apps as a source of information [13,14]. In addition, pregnant women, particularly in developed nations, are using social media to search for information about pregnancy, birth, and breastfeeding [15]. Women conduct Web-based research regularly, most often in the early stages of pregnancy when they have recently entered into a new and possibly frightening life situation [16]. Over and above other major benefits such as anonymity, simplicity, and rapidity, the major reason for searching the Web is the explicit need for advanced knowledge on a wide range of pregnancy-related topics [17]. Moreover, pregnant women are sharing their experiences and knowledge through online communities with other mothers-to-be [17,18]. Interaction in online discussion forums influences maternal health literacy through increased awareness of health promotion and health-related knowledge. For some, the information provided by other pregnant women is valued more highly than advice from health care professionals [19], thus underlining the need for high-quality and evidence-based health information on the Web.

Recent studies on the use of eHealth during pregnancy showed that most participants trusted Web-based information. However, the major proportion of these users had none to little knowledge of websites run by nonprofit organizations [13,16]. Bernhardt et al showed that mothers of young children mainly accessed commercial websites for health information, but at the same time expressed disdain for commercial websites [20].

A closer look at the increasing use of mobile pregnancy apps presents an ambiguous picture: app developers combine claims of evidence-based expertise with attempts to engage patients as part of their promotional efforts [21,22]. Yet, many users do not actively assess content validity or consider privacy issues regarding personal data collected by these apps. It is undeniable that a significant proportion of medical websites and the majority of mHealth apps are not transparent regarding information sources and privacy policy [23]. It is important, therefore, that health care professionals and pregnant women are aware of these influencing factors, their benefits, and their limitations [24]. Major providers such as Apple recently took steps to tighten requirements for medical apps (ie, more stringent data protection regulations) [25].

Notwithstanding the potentially negative connotations, the use of Web and mHealth apps undoubtedly opens new spaces for future progress. Several studies focusing on the use of eHealth and mHealth during the prenatal period demonstrated increased patient satisfaction and engagement regarding weight and blood pressure control during pregnancy [11,26-28], breastfeeding over time [29], and enhancement of subjective well-being and self-management [10,30] through the use of digital monitoring systems.

Whereas the use, the user characteristics, and the content of pregnancy-related apps are well described [14,17,31,32], qualitative studies focusing on the patients' perspective are sparse. Knight-Agarwar et al described the process of developing and pilot testing a mobile app to monitor gestational weight gain. Participants found the app motivational and approved of the nutritional information, but criticized usability [33]. Two very recent Australian studies investigated the use and value of digital media and pregnancy apps, including attitudes toward the information provided, required features (apps as tracking device or photo storage), and reservations on data protection [22,34]. The authors found that pregnant women placed high value on the information and support received through Web-based sources and apps; yet, considerations on content validity or issues concerning data security varied.

Despite the increasing influence of the Web and the rapid evolution of mHealth technology, little has changed in the prenatal care visit structure over the past years. In addition, we still do not know how pregnant women engage with mobile technology, how these mobile tools affect medical care, and whether the apparent benefits they promise are provided [21,35,36].

To fill this gap, this trial aimed to analyze (1) perceptions and requirements for pregnancy apps from the patients' perspective and (2) their impact on daily clinical routine by using qualitative research methods. We specifically decided to use a qualitative approach to obtain an unfiltered, unbiased impression of what really matters to pregnant women. In particular, we focused on the question of what features such apps must provide to meet the patients' expectations, and thus improve pregnancy care.

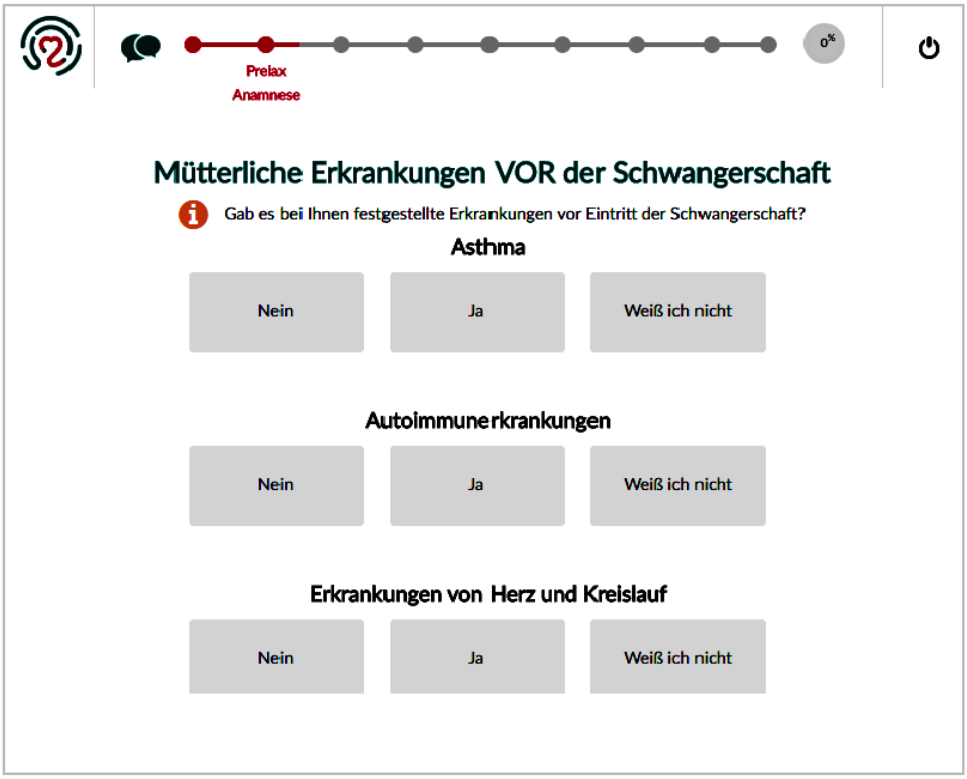
Methods

Sample

A mixed-method study with quantitative and qualitative approaches was carried out among pregnant women who attended prenatal care at the university hospital of Heidelberg from May to July 2016, a perinatal center of the highest level providing health services to low, medium, and high-risk obstetrical patients, and performing over 2000 deliveries per year. Criteria for eligibility included age of 18 years or older and a sufficient knowledge of the German language. In total, 37 randomly selected pregnant women were asked to participate,

30 of whom agreed. Reasons for not participating included lack of time or interest or poor physical condition. The sample consisted of healthy and low- to medium-risk patients to detect different user preferences and needs. Half of the group were seen during outpatient care (n=15) and half were hospitalized for the risk of preterm birth (n=15). No acute or high-risk patients were included in the study. No compensation was offered to the participants. The participants completed self-administered questionnaires on medical data, sociodemographics, and private use of technologies, developed and validated by an expert panel of doctors and midwives. Semistructured interviews were carried out with each patient to gain detailed insight in user perceptions of eHealth solutions in general after demonstrating the purpose of the questions based on a very basic patient engagement pregnancy application (PRELAX). This Web-based application consisted of a simple survey tool comprising several digital questionnaires and an information tool covering important evidence-based, pregnancy-related topics, such as nutrition, sports, body care, and nursing (Figures 1 and 2). Ethics approval was granted by the ethical committee of the University of Heidelberg.

Figure 1. Electronic data capture (eDC) tool for patient reported outcome (PRO) data.

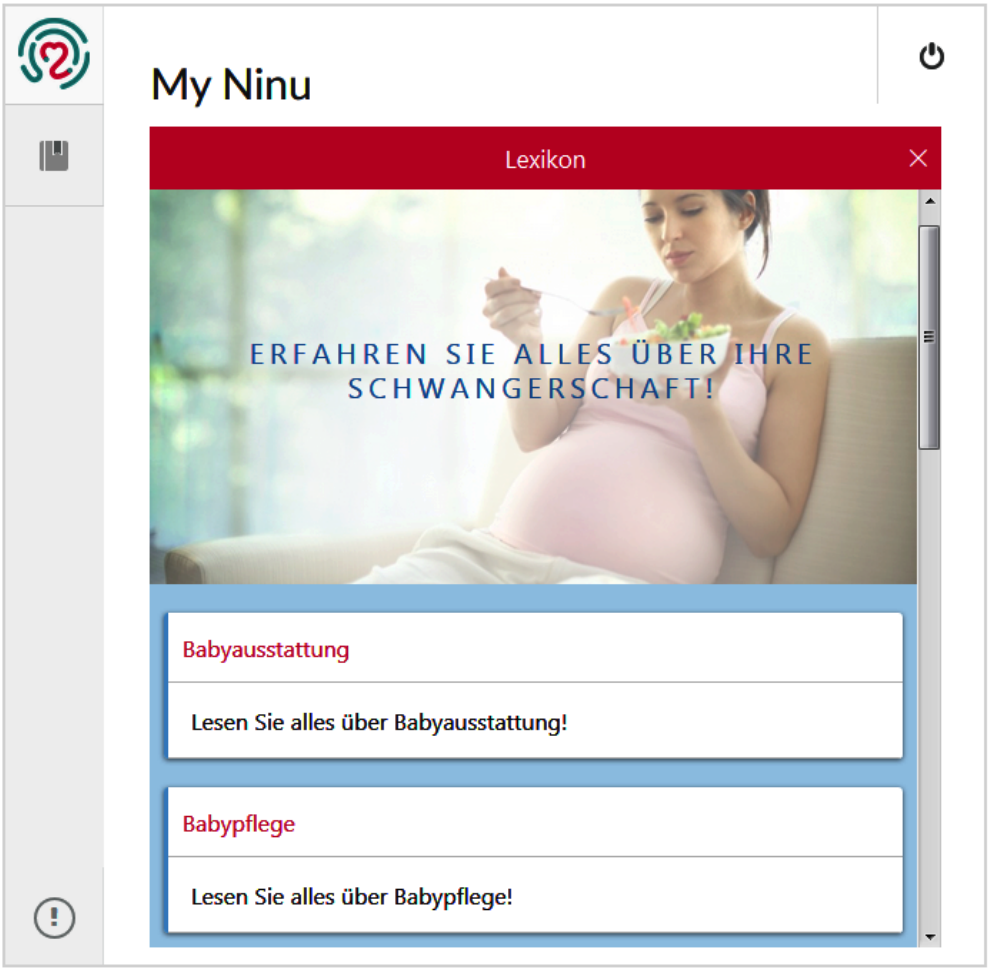


Measurements

We gathered medical and sociodemographic data such as marital status, education, and level of employment. Medical data included current diagnoses, previous pregnancies, miscarriages, and number of live-births and were double-checked against the hospital records. The private use of technologies was assessed based on individual use duration and frequency. Prior usage of the Web for information acquisition and interest in pregnancy eHealth applications or mobile apps were also assessed by the question: "Have you ever searched the Internet or a mobile app for specific questions concerning your pregnancy/hospital/clinical examinations?" Additionally, the participants were asked to complete a questionnaire focusing on the technical and graphical implementation and user-friendliness of pregnancy applications in general. The acceptance of introducing eHealth applications in clinical health care and the conceivable impact on the patients' health status, treatment, and quality of life was also assessed.

Apart from quantitative measurements, we used qualitative methods to gain a deeper understanding of the patients' view through first-hand experience and truthful quotations of semistructured face-to-face interviews. Those interviews were carried out by a trained interviewer under the supervision of a senior physician with expertise in the field of perinatology and prenatal diagnostics. The participants were aware that the interviewer was connected with the hospital but was at no time involved in their personal health care. Moreover, women were reassured about confidentiality and encouraged to openly criticize and provide lively feedback on the care they had received. After giving written consent, the interviews were digitally recorded and lasted 37 min on average (34-43 min). The interview guideline consisted of open-ended questions to assess (1) user satisfaction with the patient engagement pregnancy application developed and (2) usability and personal requirements for the use of eHealth and mHealth services (see Textbox 1 "interview section").

Figure 2. Patient education application “lexicon”.



The interviewer was free to follow spontaneous lines of thought through a flexible use of the interview schedule. Finally, the interviews were transcribed and all personal, identifying data removed.

Data Analysis

Audio-recorded interviews were transcribed verbatim and verified against the actual recordings by the authors. The authors reviewed these audio recordings for completeness of captured data in the notes. Transcripts were analyzed using systematic qualitative thematic analysis [37]. First, this approach began with familiarization with the data. Key issues and initial codes were identified, edited, and grouped into emerging themes to form the basis of the coding framework. In relation to the original data, these codes and themes were refined to ensure theoretical connectedness [38] and were finally defined. The study team discussed each stage of the analysis process to ensure correctness of the themes and their supporting data. We used interrater reliability, constant data comparison, and proper audit trail of material and processes as validation strategies. Quotes that reflected the various findings from the original data were collected for the manuscript (abbreviations: patient [P]; ambulatory [A], stationary [S]) and translated into English. Qualitative analysis software (QDA Miner Lite 1.4.5) was used to facilitate data organization using coding frequency, retrieval, and filter functions.

Textbox 1. Interview section.

Questions used to guide the interviews

- Do you think the digital survey was adequately tailored to your specific situation or would you rather have different questions?
- Considering the graphic design of the application presented: what did you like, what would you like to change?
- What should pregnancy applications provide to meet pregnant women’s expectations and to improve pregnancy care?
- What kind of incentives must be offered to pregnant women to use Web-based or mobile pregnancy applications regularly?
- What do you think about using applications surveying your medical care, health status, or quality of life (ie, mobile apps, Web-based portals) regularly in clinical routine?

Results

Demographics

In all, 30 pregnant women were included in the final study sample. Mean (standard deviation [SD]; range) maternal age was 33 years (3.4; 27-40) and mean gestational age was 33 weeks (4.3; 25-40). Demographic and pregnancy-related characteristics of the study population are shown in Table 1. The most common diagnoses of the participants are listed in Table 2.

Regarding information procurement, 26 women (87%, 26/30) frequently used Web sources to gather information on pregnancy-related topics whereas 18 women (60%, 18/30) additionally used a mobile pregnancy app. Furthermore, 24 women sought information on prenatal checkups (80%, 24/30) and 19 participants searched the Web for local maternity hospitals (63%, 19/30). Only 5 women used online communities regularly. In all, 87% (26/30) showed interest in patient engagement applications and were willing to complete health surveys electronically or via mobile devices in the future. Misgivings about the use of digital surveys in daily routine care included preference of personal treatment (n=4), concerns regarding data security (n=3), and others (n=1). A total of 22 women believed that introducing digital health surveys would improve clinical health care (73%, 22/30), 2 women assumed possible deterioration, and 6 women expected no effect at all (20%, 6/30).

Interview Findings

After analyzing data, three key themes emerged: (1) demanding expectations and perceptions for Web-based pregnancy applications, (2) favorable impact on doctor-patient relationship, and (3) frequent use and challenging requirements of eHealth applications during pregnancy. To quantify these findings, the number of counts and the coding frequency are indicated in brackets whenever possible and appropriate.

Table 1. Sample characteristics (N=30).

Characteristics	Frequency	n (%)
Education level		
Low secondary education	1	1 (3)
High secondary education	8	8 (27)
Advanced technical college entrance qualification	5	5 (17)
University entrance qualification	16	16 (53)
Total	30	30 (100)
Social status		
Married and living together	22	22 (73)
Married but living apart	1	1 (3)
Single	6	6 (20)
Divorced	1	1 (3)
Total	30	30 (100)
Level of employment		
Full-time	17	17 (57)
Part-time	9	9 (30)
Temporary exempted	4	4 (13)
Total	30	30 (100)
Current professional position		
Employee	25	25 (83)
Civil servant	5	5 (17)
Total	30	30 (100)
Gravidity		
First pregnancy	10	10 (33)
Second	11	11 (37)
Third or more	9	9 (30)
Total	30	30 (100)

Table 2. Diagnoses (N=30; multiple answers possible).

Diagnoses	Frequency	n (%)
Uneventful pregnancy	3	3 (10)
Cervical insufficiency	11	11 (37)
Preterm labor	6	6 (20)
Gestational diabetes	4	4 (13)
Placenta previa	3	3 (10)
Status after Caesarean section	3	3 (10)
Preeclampsia	3	3 (10)
Vaginal bleeding	2	2 (7)
Others	7	7 (23)

Expectations and Perceptions for Web-Based Pregnancy Applications

Evidence-Based Information

Most pregnant women criticized the lack of scientifically validated Web sources about relevant topics such as fetal development, nutrition, or pregnancy-related complications. From the patients' perspective, reliable information such as patient centered medical guidelines should be available on the Web to prevent uncertainty.

> *...because there is so much information out there, and when you get some sort of guideline, just like yours (relating to the PRELAX app.) Then that's something different from reading magazines or other Web-based portals.* [S11]

Health education also played a major role for many participants (16/30; 53%), particularly for women with a low educational level or missing medical background.

> *(It would be important) that the patients were more informed...Because I have the feeling that many pregnant women, especially the less educated, just know too little about their pregnancies...Maybe through answering the (survey) questions the women might experience a kind of wow-effect.* [A14]

The demand for information was high, in both the inpatient and outpatient setting. The need for solid background information through pregnancy Web-based applications and mobile apps was of major priority.

Personalization

Furthermore, the participants missed a certain personal touch in most Web-based applications (11/30; 37%). A personal welcoming message and the possibility to upload pictures such as fetal ultrasound images were attractive for many women.

> *Or perhaps...adding a little personal touch. That you do the log-in and then you read: "Hello, nice to have you back, Mrs. Smith". And, I don't know..., maybe also a profile and pictures, something like that...Then it would be more individual and thus more interesting for the women...* [A10]

For some participants it was difficult to cope with unexpected pregnancy complications such as miscarriages or preterm labor and many would have liked to share as much information as possible.

> *You would like to transmit personal information, indeed; especially...issues that weigh heavily on your mind...So that you have a stronger personal link to it.* [S2]

Overall, many participants assumed that the willingness to use Web-based applications regularly would increase significantly, if they were more related to the users' personal concerns.

The Role of Artificial Intelligence and Individualized Feedback Algorithms

Additionally, more than one-third of the pregnant women (11/30; 37%) expected individualized feedback algorithms via artificial intelligence software on the data they entered, either as a score or an interactive graphic. Furthermore, many would have liked to receive immediate replies to conspicuous answers and practical advice on how to proceed when experiencing specific conditions.

> *That one could simply say: "Oh, the patient's blood pressure is relatively high" and then, I don't know, just a few simple measures, how to lower it...This would add tremendous value for most women.* [A10]

Interaction: Community-Based Features

Most pregnant women found features to interact and exchange views with other mothers very useful.

> *So, what I've noticed among my friends: usually women want to share experiences, make comparisons...Just like: "Well, the doctor has told me recently that my baby weighs about 2.5 kg in the ultrasound measurement. Is this actually normal?* [A10]

Therefore, a common request was standard integration of communication platforms in pregnancy applications. In contrast to already existing online pregnancy forums, these communication tools should be open for lay people and medical professionals who could then be contacted directly via a chat feature.

> *Somehow there should be a chat room, where a doctor is available or other medical staff. So, one could ask...Specific questions...and you might even get a response from a professional.* [A8]

Usability Requirements

As many participants were frequent mobile phone (25/30; 83%) and tablet (14/30; 47%) users, usability requirements were consistently high. While pilot testing the new PRELAX application, most women criticized nonfunctional buttons or slow loading of a page (16/30; 53%).

> *The pages must load within seconds. Otherwise it's obviously really annoying...When the pages don't load quickly, then you might eventually get tired.* [S11]

In general, all women preferred an easy-to-use interface in Web-based applications and did not want to be held up with time-consuming technical issues.

Impact on Doctor-Patient Relationship

Before exploring the influence of using eHealth solutions on the doctor-patient interaction, the participants were asked for their perceptions of the current clinical routine.

In the experience of many women (9/30; 30%), doctors did not take enough time for direct doctor-patient interaction; especially time spent on rounds or conversations was considered far too short. In addition, some women felt that their physicians did not get involved enough with their issue and the average self-perceived knowledge on the patient records was rated as insufficient.

> *At the beginning it was a bit tricky here: there were always a lot of different doctors who then said: "Hum, sorry, but what exactly did you have?" It was a bit exhausting for me to explain this again and again, and then: "Oh, sure, that's what you had!"* [S12]

Another repeatedly mentioned concern (8/30; 27%) was that women could not discuss their questions due to time-pressure and their worries about clinical examinations.

> *Usually the consultation goes by really quickly. Most of the time you are so excited and nervous thinking: "How's my little baby?" that, after you have left the hospital, you notice: "Oh, I forgot to ask this or that once again."* [S14]

To solve this issue, several women (16/30; 53%) suggested using medical applications or Web portals to be prepared for the consultation. A printed medical report summarizing the most important facts could then guide the patients through the conversation with their physician.

> *For example, if you use medically-validated applications right before the medical consultation, then I think it's great; because you simply have a good basis for what you want to discuss with your doctors. During the conversation you probably won't forget that much then.* [A10]

Interpersonal relations and face-to-face conversations played an outstanding role in the doctor-patient relationship, especially for many inpatients (12/30; 40%). Several participants (8/30; 27%) stated that a Web-based pregnancy application, however good it may be, could never replace individual medical care.

> *...what I have also noticed in those apps: if you are asked...to say how you feel...That's always just a matter of interpretation, for example what "good" really means to you. This is an initial assessment to know where you stand. But it definitely cannot replace a private conversation—for me, this is still extremely important.* [A7]

However, the participants approved of integrating new technologies in clinical routine as an excellent addition (7/30; 23%).

> *I like this principle because...I know exactly, that via tablet one would admit things you wouldn't necessarily tell the doctor or nurse. So, for starters, you can state it in the application. Of course, a conversation shouldn't be missed afterwards, but this might make it easier for you to overcome yourself.* [S14]

The majority of women (26/30; 87%) wanted applications to be implemented in routine pregnancy care in order to detect and prevent serious pregnancy conditions already at an early stage.

Use and Requirements of eHealth Applications During Pregnancy

Information Procurement

For most women (26/30, 87%), the Web was considered the major source of information, providing information quickly and easily. However, some participants criticized the deficient quality of Web resources, in particular online communities for future mothers. Several women pointed out that after having read forum entries they felt insecure and confused.

> *If you have a certain problem, you quickly start reading these online forums...I think sometimes this is not so good, because what you read can be a bit unsettling." Interviewer: "Is it also important for you to know who provides the information?" P: "Actually yes!...If you hear the same thing from a doctor, it's quite different to random women writing about their stuff (laughs)."* [A11]

Thus, some participants used Google to look up pregnancy symptoms they were experiencing, but would not recommend it.

Several participants were already using mobile pregnancy apps regularly (18/30; 60%). They explicitly liked to retrieve suitable information about relevant pregnancy topics on a weekly basis, hence, not too much information at once. For many women, mobile phone apps were considered to be playful but also useful tools.

> *I think pregnant women are more likely to use a mobile app..., which assesses various issues weekly. So, you know, one could feel somehow accompanied medically. But it wouldn't be too much information*

> *all at once. I think those technical solutions would be attractive for pregnant women.* [S2]

Attitude Toward Digitization

In general, many participants (18/30; 60%) had a positive attitude toward the growing eHealth movement and could imagine a more substantial role for applications in daily routine. For example, one woman suggested connecting mobile lifestyle apps with medical applications to gather as much data as possible.

> *I really think this is appropriate for our age. I mean, if you look around these days, there is an app for everything!!...And that is also the trend for the future. Especially in the medical fields such tools would be great with already existing standard lifestyle apps. It would be nice, if you combined these maybe. Then you would have additional parameters or data on the general health, sleep patterns, nutrition...* [S10]

Furthermore, sharing data among health care professionals (ie, physicians and midwives) or even health insurance companies through standardized networks constituted a major request (7/30; 23%).

> *What you mentioned earlier: that you could forward your data directly to your gynecologist, midwife, hospital etc., and they equally have access to the patient's data, for example. In a digital age like ours that would probably be very useful. Then you wouldn't have to fill out another form each time.* [A7]

Some women also hoped that the technical advantages might enable faster interventions in case of any critical pregnancy condition. Although most participants appreciated the implementation of applications in pregnancy care because of their easy-use and ubiquitous availability, there was disagreement on the handling and transfer of personal data.

Data Security

Notably, hospitalized women (6/30; 20%) were worried about unauthorized third-party accesses to their stored medical data.

> *Of course, on the other hand one is always afraid that the personal data might go anywhere, and there is also the risk of unauthorized accesses to your data you actually do not agree with.* [S5]

Various women (8/30; 27%) expressed concerns about data security, especially in the field of mobile apps, since many free apps make private data easily accessible.

> *As I said, I'm very critical about patient data in general, especially in terms of data security...If you have a free app, it really depends on what happens to the private data. As a matter of fact, usually the information is stored on the app itself, and so other apps might gain access to the data easily.* [A5]

Despite these misgivings, several women were willing to transmit their personal data at the touch of a button.

> *Of course...it would seem reasonable that a considerable amount of data could be forwarded to different recipients by using such tools. Also, my attitude towards data security isn't like: "For God's sake: Cannot be, must not be!" Because, I think that only those will have access to the data who should have. So, I do believe that such things might facilitate certain procedures a lot.* [A12]

Personal Incentive

For many participants (11/30; 37%), there was no need to offer special incentives to use eHealth applications. However, some women pointed out that they would not have used an application surveying their pregnancy course from home if they had not shown any symptoms.

> *Personally I would be satisfied, if I could obtain information about my particular situation...But, to be honest, if I were completely healthy and had an uncomplicated pregnancy from the beginning to the end, then my incentives here would be rather low. Then, I would probably need something else, maybe like individual health care procedures, that kind of thing.* [A6]

In conclusion, most women (20/30; 70%) wished to benefit from the new technological solutions and take best advantage of them for the course of their pregnancy. Particularly the outpatient group stated that using such applications wasn't seen as an end in itself and they were willing to contribute to implementing pioneering eHealth applications (6/30; 20%).

Discussion

Principal Findings

In our study, the following main requirements of eHealth applications for pregnancy emerged (see Textbox 2 "key findings of the interviews"): Most pregnant women, regardless of age and health status, criticized the poor quality of existing Web-based information sources and mobile apps. They had an obvious need for scientifically validated information about pregnancy-related topics. Regarding the expectations for patient engagement applications, our findings revealed a strong request for individualized feedback algorithms and individually tailored information. Several women favored more interactive apps and recommended communication platforms for both pregnant women and medical professionals. In general, usability requirements were high and the women stressed the need for a user-friendly interface in Web-based applications and mobile apps. Since many participants experienced a lack of time in the doctor-patient interaction, some suggested using Web-based applications to be better prepared for the consultation. The majority of the participants also approved of integrating digital media and modern technological devices in clinical routine and pregnancy care due to their easy-use and ubiquitous availability.

Nevertheless, several concerns emerged: a considerable number of respondents had reservations, especially concerning the safety and storage of personal data in electronic databases or applications.

Overall, no significant difference could be detected regarding user behavior and requirements in the 2 sample groups—the inpatient or outpatient participants. This could indicate that the

use of eHealth or mHealth doesn't primarily depend on the user's health status, but instead on the socioeconomic background or education. Further research with a wider socioeconomic range in the study population is required to identify different needs.

Textbox 2. Key findings of the interviews.

Key findings

- Mobile and Web-based pregnancy applications as a highly frequented source for evidence-based information about most relevant pregnancy topics
- Strong request for a more personalized output and preference for interactive applications (eg, individualized feedback algorithms, community-based features)
- Impact on doctor-patient relationship: fostering patient empowerment and a partner-like relationship
- Openness for integrating eHealth or mHealth applications in daily pregnancy care and potential digital networking among health care providers
- Data security and personal data storage in pregnancy applications as general cause for concern

Comparison With Prior Work

Our findings are in line with other recent trials. Several studies showed that pregnant women use the Web and digital media to improve their knowledge [16,20,26,36,39] and how they evaluate evidence-based information in pregnancy applications [17]. However, previous research has reported diverging findings on the reliability of Web resources and confidence in information offered by medical professionals. Whereas Bert et al reported that 70% percent of the study participants referred to the Web as a highly reliable information source [17], our sample had a more critical stance on this. Inter alia, Kraschnewski et al found that many future mothers asked "Dr Google" first when experiencing unknown pregnancy symptoms [36]. Most of our participants first "googled" pregnancy-related symptoms, but did not experience positive reinforcement; hence, they would not recommend it.

Most of the women interviewed demanded rapid and easy access to evidence-based content on digital media. Lupton et al showed that information offered by professionals was highly valued when women had a specific, health-related concern [34]. In our sample, the need for such information was high in both inpatient and outpatient settings. Since all participants chose to deliver in a level-one prenatal center, the general concern regarding their pregnancy and their child's health was higher.

Among others, studies from Australia and Norway reported that many pregnant women enjoyed the emotional support of other mothers through social media [19,34]. However, most of our study participants expressed reservations about online communities as a source of erroneous information. In accordance with Fredriksen et al, this may be due to the fact that the level of education in our sample was higher than average.

Relating to data security, our findings suggest that hospitalized women in particular were aware of the potential harms of applications with low data protection standards. Then again, various participants raised little concern about personal data security, which is consistent with other findings [17,22,40].

In line with other studies [22,24,34], the women in our sample appreciated interactive modules in applications. The emphasis on a personalized output of the apps could be part of the demand for a more personal support, feeling somehow they have medical support through patient-tailored applications. Lee et al showed that whereas a social networking function was important for pregnant women, interaction with health professionals still remained limited [32,36]. Standard integration of communication platforms open for lay persons and experts in such applications would provide the users with professional and instantaneous advice, leading to a more partner-like relationship.

Whereas Tripp et al assumed that reliance on health care professionals might be reduced by interactive and personalized information delivered via mobile phones [24], Ledford et al showed that no difference was detected on interpersonal clinical communication [10]. Yet, our observations showed a different result, as most women in our sample stressed the importance of individual patient-centered care. eHealth applications were seen as helpful, but more as complementary tools.

Strengths and Limitations

As far as we know, this is the first qualitative study to analyze perceptions and expectations of pregnancy applications from the patients' perspective in Germany, thus creating a basis for further research. This study contained a small sample of n=30, which may be considered as a limitation. However, this sample size is generally accepted for qualitative studies. Apart from ensuring richness of data through a qualitative design, the study demonstrated theoretical connectedness by using direct quotes from the original data to support themes. Although one should take care not to generalize these qualitative findings, most results were consistent with various international qualitative [19,30,36,41] and quantitative studies [14,31,34] in this field.

Nevertheless, some limitations should be taken into account. All patients were recruited at a university hospital; therefore, more patients than average had a history of two or more miscarriages (4/30, 13%). Furthermore, the average educational level was higher, potentially resulting in greater health. This needs to be considered while comparing our results to average community hospital populations. Further research among socioeconomically disadvantaged women with a lower educational level is required to identify different needs.

Conclusions

Whereas previous research has explored the content and use of pregnancy-related applications, our study provides insight into

the patients' perceptions and expectations. We showed that pregnant women considered evidence-based information and interactive tools as the most important features. Therefore, developing medically accurate eHealth and mHealth applications poses a challenge to interdisciplinary app developers.

The next evolutionary step is to successfully integrate these evidence-based medical applications into daily health care practice, fostering both patient engagement and empowerment. Health care professionals should be committed to guiding pregnant women through these applications, exploring the ability to prevent misguiding through nonvalidated educational information, and thus reduce adverse pregnancy outcomes.

Conflicts of Interest

None declared.

References

1. Bartsch P, Lux T, Wagner A, Gabriel R. Business intelligence and information systems in hospitals--distribution and usage of BI and HIS in German hospitals. Stud Health Technol Inform 2013;190:191-193. [Medline: 23823419]
2. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun;3(2):e20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
3. Eysenbach G, Diepgen TL. The role of e-health and consumer health informatics for evidence-based patient choice in the 21st century. Clin Dermatol 2001;19(1):11-17. [Medline: 11369478]
4. Davis TL, DiClemente R, Prietula M. Taking mHealth forward: examining the core characteristics. JMIR Mhealth Uhealth 2016 Aug 10;4(3):e97 [FREE Full text] [doi: 10.2196/mhealth.5659] [Medline: 27511612]
5. World Health Organization (WHO). mHealth: New horizons for health through mobile technologies. In: Global Observatory for eHealth series - Volume 3. Geneva, Switzerland: World Health Organization; 2011.
6. Trotter MI, Morgan DW. Patients' use of the Internet for health related matters: a study of internet usage in 2000 and 2006. Health Informatics J 2008 Sep;14(3):175-181. [doi: 10.1177/1081180X08092828] [Medline: 18775824]
7. Grosberg D, Grinvald H, Reuveni H, Magnezi R. Frequent surfing on social health networks is associated with increased knowledge and patient health activation. J Med Internet Res 2016 Aug 10;18(8):e212 [FREE Full text] [doi: 10.2196/jmir.5832] [Medline: 27511272]
8. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. J Med Internet Res 2013;15(2):e37 [FREE Full text] [doi: 10.2196/jmir.2281] [Medline: 23425685]
9. Berrewaerts J, Delbecque L, Orban P, Desseilles M. Patient participation and the use of ehealth tools for pharmacoviligance. Front Pharmacol 2016;7:90 [FREE Full text] [doi: 10.3389/fphar.2016.00090] [Medline: 27148052]
10. Ledford CJ, Canzona MR, Cafferty LA, Hodge JA. Mobile application as a prenatal education and engagement tool: a randomized controlled pilot. Patient Educ Couns 2016 Apr;99(4):578-582. [doi: 10.1016/j.pec.2015.11.006] [Medline: 26610389]
11. Marko KI, Ganju N, Brown J, Benham J, Gaba ND. Remote prenatal care monitoring with digital health tools can reduce visit frequency while improving satisfaction [3]. Obstet Gynecol 2016;127:1S. [doi: 10.1097/01.AOG.0000483620.40988.df]
12. 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]
13. Lima-Pereira P, Bermúdez-Tamayo C, Jasienska G. Use of the Internet as a source of health information amongst participants of antenatal classes. J Clin Nurs 2012 Feb;21(3-4):322-330. [doi: 10.1111/j.1365-2702.2011.03910.x] [Medline: 22093043]
14. Wallwiener S, Müller M, Doster A, Laserer W, Reck C, Pauluschke-Fröhlich J, et al. Pregnancy eHealth and mHealth: user proportions and characteristics of pregnant women using Web-based information sources-a cross-sectional study. Arch Gynecol Obstet 2016 Apr 15 Epub ahead of print. [doi: 10.1007/s00404-016-4093-y] [Medline: 27084763]
15. McCann AD, McCulloch JE. Establishing an online and social media presence for your IBCLC practice. J Hum Lact 2012 Nov;28(4):450-454. [doi: 10.1177/0890334412461304] [Medline: 23087194]
16. Larsson M. A descriptive study of the use of the internet by women seeking pregnancy-related information. Midwifery 2009 Feb;25(1):14-20. [doi: 10.1016/j.midw.2007.01.010] [Medline: 17408822]
17. Bert F, Gualano MR, Brusaferro S, De VE, de WC, La TG, et al. Pregnancy e-health: a multicenter Italian cross-sectional study on Internet use and decision-making among pregnant women. J Epidemiol Community Health 2013 Dec 1;67(12):1013-1018. [doi: 10.1136/jech-2013-202584] [Medline: 24072743]
18. Johnson SA. 'Intimate mothering publics': comparing face-to-face support groups and Internet use for women seeking information and advice in the transition to first-time motherhood. Cult Health Sex 2015;17(2):237-251. [doi: 10.1080/13691058.2014.968807] [Medline: 25339096]
19. Fredriksen EH, Harris J, Moland KM. Web-based discussion forums on pregnancy complaints and maternal health literacy in Norway: a qualitative study. J Med Internet Res 2016 May 26;18(5):e113 [FREE Full text] [doi: 10.2196/jmir.5270] [Medline: 27230094]

20. Bernhardt JM, Felter EM. Online pediatric information seeking among mothers of young children: results from a qualitative study using focus groups. J Med Internet Res 2004 Mar 1;6(1):e7 [FREE Full text] [doi: 10.2196/jmir.6.1.e7] [Medline: 15111273]
21. Lupton D, Jutel A. 'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps. Soc Sci Med 2015 May;133:128-135. [doi: 10.1016/j.socscimed.2015.04.004] [Medline: 25864149]
22. 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]
23. Bert F, Giacometti M, Gualano MR, Siliquini R. Smartphones and health promotion: a review of the evidence. J Med Syst 2014 Jan;38(1):9995. [doi: 10.1007/s10916-013-9995-7] [Medline: 24346929]
24. 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]
25. Apple. 2016. App store review guidelines URL: https://developer.apple.com/app-store/review/guidelines/ [accessed 2016-11-06] [WebCite Cache ID 6loRHEwu3]
26. Waring ME, Moore ST, Xiao RS, Lombardini LM, Allison JJ, Rosal MC, et al. Pregnant women's interest in a website or mobile application for healthy gestational weight gain. Sex Reprod Healthc 2014 Dec;5(4):182-184. [doi: 10.1016/j.srhc.2014.05.002] [Medline: 25433828]
27. Demment MM, Graham ML, Olson CM. How an online intervention to prevent excessive gestational weight gain is used and by whom: a randomized controlled process evaluation. J Med Internet Res 2014;16(8):e194 [FREE Full text] [doi: 10.2196/jmir.3483] [Medline: 25143156]
28. Marko KI, Krapf JM, Meltzer AC, Oh J, Ganju N, Martinez AG, et al. Testing the feasibility of remote patient monitoring in prenatal care using a mobile app and connected devices: a prospective observational trial. JMIR Res Protoc 2016 Nov 18;5(4):e200 [FREE Full text] [doi: 10.2196/resprot.6167] [Medline: 27864167]
29. O'Connor ME, Brown EW, Lewin LO. An Internet-based education program improves breastfeeding knowledge of maternal-child healthcare providers. Breastfeed Med 2011 Dec;6(6):421-427. [doi: 10.1089/bfm.2010.0061] [Medline: 21029021]
30. Haga SM, Drozd F, Brendryen H, Slinning K. Mamma mia: a feasibility study of a web-based intervention to reduce the risk of postpartum depression and enhance subjective well-being. JMIR Res Protoc 2013;2(2):e29 [FREE Full text] [doi: 10.2196/resprot.2659] [Medline: 23939459]
31. Bert F, Passi S, Scaioli G, Gualano MR, Siliquini R. There comes a baby! What should I do? Smartphones' pregnancy-related applications: a web-based overview. Health Informatics J 2016 Sep;22(3):608-617. [doi: 10.1177/1460458215574120] [Medline: 25900813]
32. Lee Y, Moon M. Utilization and content evaluation of mobile applications for pregnancy, birth, and child care. Healthc Inform Res 2016 Apr;22(2):73-80 [FREE Full text] [doi: 10.4258/hir.2016.22.2.73] [Medline: 27200216]
33. Knight-Agarwal C, Davis DL, Williams L, Davey R, Cox R, Clarke A. Development and pilot testing of the Eating4two mobile phone app to monitor gestational weight gain. JMIR Mhealth Uhealth 2015;3(2):e44 [FREE Full text] [doi: 10.2196/mhealth.4071] [Medline: 26048313]
34. Lupton D. The use and value of digital media for information about pregnancy and early motherhood: a focus group study. BMC Pregnancy Childbirth 2016 Jul 19;16(1):171 [FREE Full text] [doi: 10.1186/s12884-016-0971-3] [Medline: 27435182]
35. Lupton D. Health promotion in the digital era: a critical commentary. Health Promot Int 2015 Mar;30(1):174-183. [doi: 10.1093/heapro/dau091] [Medline: 25320120]
36. Kraschnewski JL, Chuang CH, Poole ES, Peyton T, Blubaugh I, Pauli J, et al. “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;16(6):e147 [FREE Full text] [doi: 10.2196/jmir.3385] [Medline: 24892583]
37. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
38. Burns N. Standards for qualitative research. Nurs Sci Q 1989;2(1):44-52. [doi: 10.1177/089431848900200112] [Medline: 2927791]
39. Lagan B, Sinclair M, Kernohan W. Internet use in pregnancy informs women's decision making: a web-based survey. Birth 2010 Jun;37(2):106-115.
40. Pandey SK, Hart JJ, Tiwary S. Women's health and the internet: understanding emerging trends and implications. Soc Sci Med 2003 Jan;56(1):179-191. [Medline: 12435560]
41. Odeny TA, Newman M, Bukusi EA, McClelland RS, Cohen CR, Camlin CS. Developing content for a mHealth intervention to promote postpartum retention in prevention of mother-to-child HIV transmission programs and early infant diagnosis of HIV: a qualitative study. PLoS One 2014;9(9):e106383 [FREE Full text] [doi: 10.1371/journal.pone.0106383] [Medline: 25181408]

Abbreviations

eDC: electronic Data Capture

PRO: Patient Reported Outcome
PRELAX: pregnancy application, term consists of “pregnancy” and “relax”

Edited by G Eysenbach; submitted 05.12.16; peer-reviewed by S Brucker, C Ledford, A Jutel; comments to author 10.02.17; revised version received 29.03.17; accepted 11.04.17; published 26.05.17.

Please cite as:
Goetz M, Müller M, Matthies LM, Hansen J, Doster A, Szabo A, Pauluschke-Fröhlich J, Abele H, Sohn C, Wallwiener M, Wallwiener S
Perceptions of Patient Engagement Applications During Pregnancy: A Qualitative Assessment of the Patient’s Perspective
JMIR Mhealth Uhealth 2017;5(5):e73
URL: http://mhealth.jmir.org/2017/5/e73/
doi:10.2196/mhealth.7040
PMID:28550005

Original Paper

A Smartphone-Based Approach for Triage of Human Papillomavirus-Positive Sub-Saharan African Women: A Prospective Study

Esther Urner[1*], MBBS; Martine Delavy[1*], MBBS; Rosa Catarino[1], MD; Manuela Viviano[1], MD; Ulrike Meyer-Hamme[1], MD; Anne-Caroline Benski[1], MD; Jeromine Jinoro[2], MD; Josea Lea Heriniainasolo[3], MD; Manuela Undurraga[1], MD; Hugo De Vuyst[4], MD; Christophe Combescure[5], PhD; Pierre Vassilakos[6], MD; Patrick Petignat[1], MD

[1]Gynecology Division, Department of Gynecology and Obstetrics, Geneva University Hospitals, Geneva, Switzerland

[2]Laboratory Division, Saint Damien Healthcare Centre, Ambanja, Madagascar

[3]Gynecology Division, Department of Gynecology and Obstetrics, Saint Damien Healthcare Centre, Ambanja, Madagascar

[4]International Agency for Research on Cancer, World Health Organization, Lyon, France

[5]Research Center, Geneva University Hospitals, Geneva, Switzerland

[6]Geneva Foundation for Medical Education and Research, Geneva, Switzerland

[*]these authors contributed equally

Corresponding Author:
Martine Delavy, MBBS
Gynecology Division
Department of Gynecology and Obstetrics
Geneva University Hospitals
Boulevard de la Cluse 30
Geneva, 1206
Switzerland
Phone: 41 22 372 68 16
Fax: 41 22 347 64 86
Email: martinegilliane@gmail.com

Abstract

Background: Sub-Saharan African countries are marked by a high incidence of cervical cancer. Madagascar ranks 11th among the countries with the highest cervical cancer incidence worldwide.

Objective: The aim of the study was to evaluate the performances of digital smartphone-based visual inspection with acetic acid (D-VIA) and Lugol's iodine (D-VILI) for diagnosing cervical precancer and cancer.

Methods: Human papillomavirus (HPV)-positive women recruited through a cervical screening campaign had D-VIA and D-VILI examinations with endocervical curettage (ECC) and cervical biopsy. Three images were captured for each woman (native, D-VIA, D-VILI) using a smartphone camera. The images were randomly coded and distributed on 2 online databases (Google Forms). The D-VIA form included native and D-VIA images, and the D-VILI form included native and D-VILI images. Pathological cases were defined as cervical intraepithelial neoplasia grade 2 or worse (CIN2+). Physicians rated the images as non-pathological or pathological. Using the ECC and cervical biopsy results as references, the sensitivity and specificity of D-VIA and D-VILI examinations for each and all physicians were calculated.

Results: Altogether, 15 clinicians assessed 240 images. Sensitivity was higher for the D-VIA interpretations (94.1%; 95% CI 81.6-98.3) than for the D-VILI interpretations (78.8%; 95% CI 54.1-92.1; *P*=.009). In contrast, the specificity was higher for the D-VILI interpretations (56.4%; 95% CI 38.3-72.9) than for the D-VIA interpretations (50.4%; 95% CI 35.9-64.8; *P*=.005).

Conclusion: Smartphone-based image for triage of HPV-positive women is more accurate for detecting CIN2+ lesions with D-VIA than D-VILI, although with a small loss of specificity.

(JMIR Mhealth Uhealth 2017;5(5):e72) doi:10.2196/mhealth.6697

KEYWORDS

cervical cancer; squamous intraepithelial lesions of the cervix; HPV; acetic acid; lugol's iodine; smartphone; mobile phone

Introduction

Sub-Saharan African countries are marked by a high incidence of cervical cancer [1]. Madagascar, in particular, counts about 3194 new cervical cancer cases every year and ranks 11th among the countries with the highest cervical cancer incidence worldwide [1].

Because of the long interval between the initial human papilloma virus (HPV) infection and the development of cervical cancer, screening has proven to be largely effective in reducing the incidence of cervical cancer in industrialized countries [2]. In low- and medium-income countries (LMICs), however, the lack of expertise and infrastructure makes screening programs difficult to implement. To overcome this difficulty, one of the options recommended by the World Health Organization (WHO) includes HPV testing followed by visual inspection of the cervix with acetic acid (VIA) as a triage test [3]. A meta-analysis, however, has shown that visual inspection with Lugol's iodine (VILI) may perform better than VIA [4].

Given the concerns about the suboptimal sensitivity of VIA/VILI tests and the lack of a quality assurance system, the use of digital VIA (D-VIA) and D-VILI is a promising choice for LMICs. Smartphone cameras are easy to use, have excellent image quality, and by sending the images to other on-site and off-site experts, they offer the possibility to obtain real-time feedback.

A previous study conducted in Madagascar has shown that off-site detection of cervical lesions based on the evaluation of smartphone photographs was more reliable than on-site diagnosis alone [5]. The aim of this study was to assess the clinical performance of D-VIA and D-VILI examinations for diagnosing cervical lesions in LMICs.

Methods

The cervical cancer screening campaign took place in the Saint Damien Healthcare Centre in Ambanja, Madagascar and in five dispensaries in the surrounding rural areas between February and October 2015. As a means of primary screening, women were asked to collect a vaginal self-sample with a sterile swab (ESwab, Copan, Brescia, Italy). The samples were then analysed in Ambanja by a point-of-care HPV assay (GeneXpert IV; Cepheid, Sunnyvale, CA, USA). Women who tested positive for high-risk HPV genotypes were invited to the Saint Damien Healthcare Centre to undergo a gynecological examination. Women were included in the study if they were aged between 30 to 69 years, HPV-positive for high-risk HPV genotypes, able to understand study procedures, and voluntarily participated by signing the informed consent form. Exclusion criteria included a previous hysterectomy, conditions that could interfere with visualization of the cervix, pregnancy over 20 weeks, or that they could not comply with study protocol. This cross-sectional analysis included data from HPV-positive women with a disease prevalence corresponding to that of real-life conditions (10–15% cervical intraepithelial neoplasia grade 2 or worse [CIN2+]).

All HPV-positive patients underwent digital VIA/VILI examinations, which were carried out on the same day or a few weeks following self-HPV testing. Three images per examination were captured using a smartphone (Samsung Galaxy S4 or S5, Seoul, South Korea). The first image corresponded to the native cervix. The second was captured one minute after application of a 5% acetic acid solution and the third following application of Lugol's iodine. Cervical (Papanicolaou) smears, endocervical curettage and an average of 2 biopsies were performed on all patients. If no lesion was identified, a random cervical biopsy was performed at the 6 o'clock site. Patients for whom a pathological cervical lesion (CIN2+) was suspected based on the immediate evaluation by the on-site expert were treated with cold coagulation or large electrosurgical excision procedure (LEEP) on the same day. A hysterectomy was scheduled for the following days when invasive cervical cancer was suspected, if operable. Women who were not promptly treated on site but who were later diagnosed histologically with cervical dysplasia were called again to the Saint Damien Healthcare Centre for further management.

The GeneXpert device was used on site for HPV DNA detection. Using a polymerase chain reaction method, this test detects 14 types of high-risk HPV. The results are divided into three sections: HPV 16, HPV 18/45, and 11 other high-risk HPV types (31, 33, 35, 39, 51, 52, 56, 58, 59, 66, and 68). Endocervical samples were collected using an endocervical brush. A cervical biopsy forceps was used for histological specimen collection. Both samples were fixed in liquid formalin. The reference standard for the disease was the histological evaluation, which was performed in Geneva, Switzerland. The results were interpreted according to the WHO 2014 classification as Low-grade squamous intraepithelial lesions (LSIL) and High-grade squamous intraepithelial lesions (HSIL) and sub-classified as grade 1, 2, or 3 CIN.

Throughout the pelvic examination, photographs were obtained at a distance of 10-15 cm of the cervix, with 3.3-3.8× optical zoom and in flash mode. The two smartphones, Samsung Galaxy S4 and S5, were chosen for their high-quality cameras (13 and 16 megapixels, respectively, both with auto-focus and flash functions). These devices allow highly precise and detailed visualization of the cervix after zooming and focusing in on the target. To improve the stability and quality of the images, the smartphone was fixed to the ground with a tripod and a support.

The images were uploaded onto two online databases (Google forms). One form included the native and D-VIA images. The second included the native and D-VILI images. The images were randomly coded and distributed on the two databases so that the D-VIA and D-VILI images of each patient could not be linked to one another. Overall, 15 physicians with different clinical backgrounds and experience, blind to the histological and HPV genotypes, were asked to determine (for each patient in each of the two Google forms) whether the images of the cervix were non-pathological (<CIN2) or pathological (CIN2+)

using IARC Reference Chart for Visual Inspection with acetic acid (VIA) and Lugol's iodine (VILI). Figures 1 and 2 represent a CIN3 case presented on the D-VIA and D-VILI forms, respectively. The physicians also completed a questionnaire addressing their own background and experience in colposcopy (Table 1). The examining physicians were labeled as "experts" if the number of colposcopies performed in their lifetime was greater than 300 and if they had at least ten years of experience in clinical practice. Out of the 15 physicians, 5 fulfilled these criteria and were labeled as "experts." The remaining 10 physicians had performed less than 300 colposcopies, had less than 10 years of experience in clinical practice, and were labeled as "novices."

The sensitivity and specificity for the detection of CIN2+ were calculated for the D-VIA and D-VILI interpretations of each rater. The correlation coefficients between screener-specific sensitivities and specificities were assessed for both D-VIA and D-VILI (Spearman coefficient of correlation). To assess the global sensitivity and specificity (for all readers), we used logistic regression models with mixed effects to account for repeated measures: a random intercept was introduced into the model, and the random effects for patients and readers were crossed. Models were applied to patients with a positive biopsy to assess sensitivity and to patients with a negative biopsy to assess specificity. Additional logistic regression models with mixed effects were used to test a difference in sensitivity or specificity between expert screeners and novice screeners and to test a difference between D-VIA and D-VILI.

All statistical tests were two-sided, with a type 1 error of 0.05. Statistical analyses were performed with R Core Team 2015 software (R Foundation for Statistical Computing, Vienna, Austria) and Stata13 IC software (StataCorp, College Station, TX, USA).

The study was approved by the Malagasy National Commission for the Ethics of Science and Technology and the Ethical Cantonal Board of Geneva, Switzerland (CER: 14–071). The trial has been registered on clinicaltrials.gov (NCT02693379).

Figure 1. Digital visual inspection using acetic acid (D-VIA). Left. Native cervix. Right. Cervix after acetic acid application.

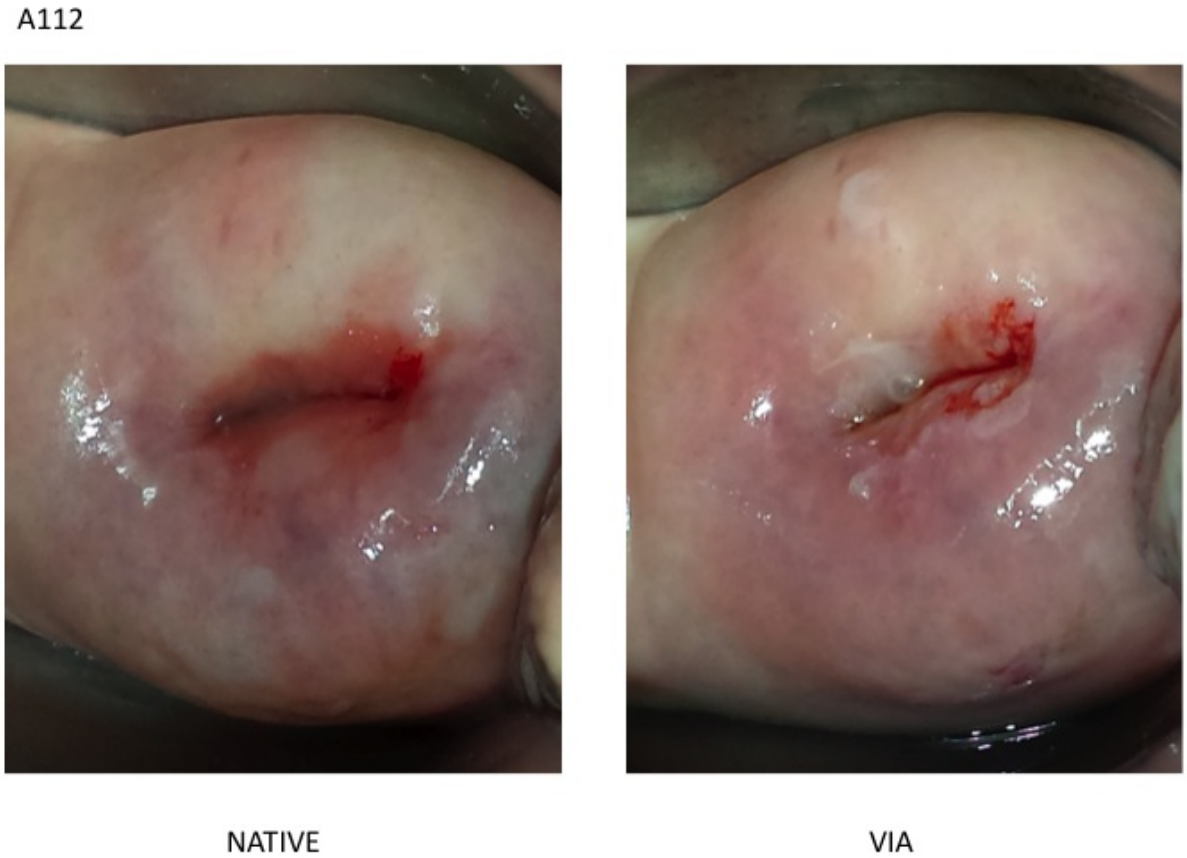
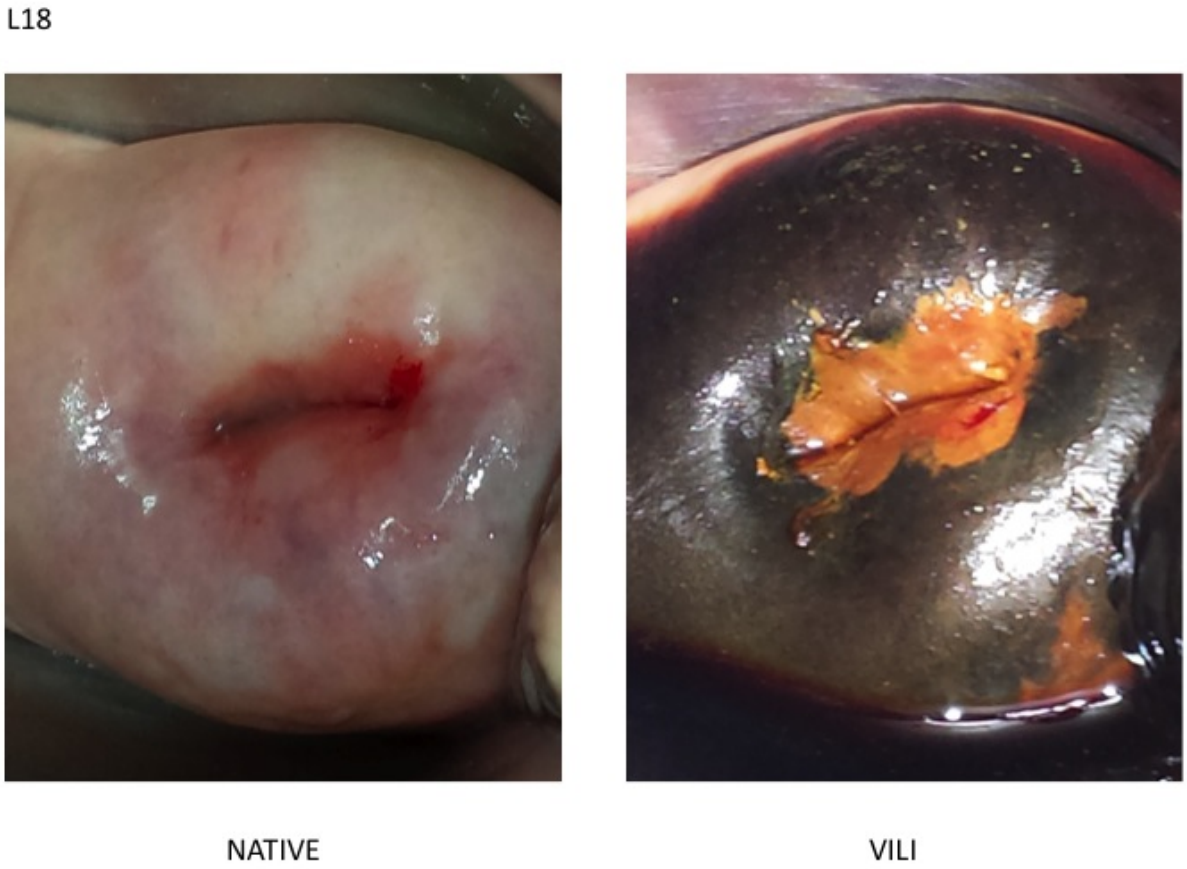


Figure 2. Digital visual inspection using Lugol's iodine (D-VILI). Left. Native cervix. Right. Cervix after Lugol's iodine application.



Results

Among the 1041 participants screened for HPV, 231 were HPV-positive (22.2%) and 584 HPV-negative (77.8%). Among HPV-positive women, 187 underwent digital VIA/VILI testing in the Saint Damien Health Care Centre. A total of 561 images were captured during the examination. A total of 120 out of 187 (64.2%) sets of photographs of women who underwent digital VIA/VILI were considered of good quality and were included in the study. Images were excluded because of insufficient focus or inadequate magnification. The resulting disease prevalence in our study corresponds to that of an unscreened asymptomatic population (10-15%) [6].

The overall distribution of cervical disease among the 120 patients was 14 (11.6%) pathological cases (CIN2+) that included 4 (3.3%) cases of invasive cancer. There were 7 (5.8%) cases of CIN3, 3 (2.5%) cases of CIN2, and 4 (3.3%) cases of CIN1. In all, 102 (85.0%) cases were negative for any cervical lesion. CIN1 lesions were considered and managed as negative cases. Overall, 14 out of 65 (21.5%) CIN2+ cases were promptly treated with thermocoagulation or conisation, whereas 51 out of 64 (79.7%) cases were overtreated.

The raters' mean age was 39.7 years (SD 8.3). Their mean number of years of experience in clinical practice was 11.0 years (SD 6.5). Table 1 reports the sociodemographic characteristics of the 15 raters.

Table 1. Raters' sociodemographic characteristics.

Variable		n (%)
Total		15
Age (years), mean (SD)		39.7 (SD 8.3)
Nationality		
	Swiss	7 (46.7)
	Other	8 (53.3)
Profession		
	Medical Doctor	14 (93.3)
	Other	1 (6.7)
Years of experience in clinical practice, mean (SD)		11.0 (SD 6.5)
Number of colposcopies during lifetime		
	0	1 (6.7)
	<50	6 (40.0)
	51–300	3 (20.0)
	>300	5 (33.3)
Number of colposcopies per year		
	0	4 (26.7)
	<30	8 (53.3)
	>60	3 (20.0)

The diagnostic performance of D-VIA and D-VILI varied broadly among raters: the overall sensitivity was 81.6-98.3% for D-VIA and 54.1-92.1% for D-VILI. The specificity was 35.9-64.8% for D-VIA and 38.3-72.9% for D-VILI. For both techniques, the specificity decreased as the sensitivity increased ($\rho=-0.59$ for D-VIA and $\rho=-0.80$ for D-VILI). The overall sensitivity was 94.1% (95% CI 81.6-98.3) for D-VIA and 78.8% (95% CI 54.1-92.1) for D-VILI (*P*=.009). The overall specificity was 50.4% (95% CI 35.9-64.8) for D-VIA and 56.4% (95% CI 38.3-72.9) for D-VILI (*P*=.005).

The specificity was higher for experts than novices, but the difference was statistically significant only for D-VIA (*P*=.04). No statistically significant differences between experts and novices were found concerning sensitivity. Among the expert raters, D-VIA had higher sensitivity than D-VILI (*P*=.01) and lower specificity (*P*=.02). In novice raters, the trend of the findings was the same, but the difference in specificities and sensitivities were not statistically significant (*P*=.07 and *P*=.05, respectively). Table 2 and Figure 3 report the sensitivity, specificity, and *P* values for D-VIA and D-VILI among all raters, experts, and novices.

Table 2. Sensitivity, specificity, and *P* value for D-VIA and D-VILI among all raters, experts, and novices

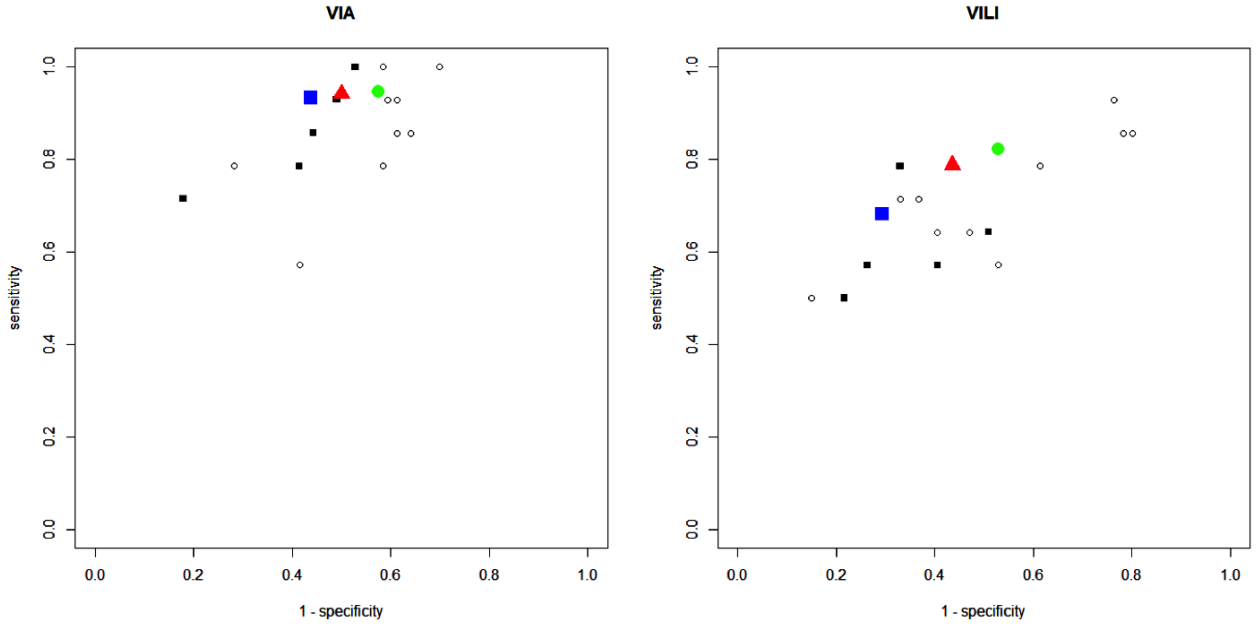
Parameters		D-VIA[a] Overall (range)	D-VILI[b] Overall (range)	*P* [c]
Sensitivity				
	All raters (n=15)	94.1% (81.6-98.3)	78.8% (54.1-92.1)	.009
	Experts (n=5)	93.3% (67.3-99.0)	68.5% (37.8-88.6)	.01
	Novices (n=10)	94.6% (78.8-98.8)	82.3% (58.3-94.0)	.05
P (experts vs novices)		.86	.12	
Specificity				
	All raters (n=15)	50.4% (35.9-64.8)	56.4% (38.3-72.9)	.005
	Experts (n=5)	56.2% (44.5-81.4)	70.6% (57.0-81.3)	.02
	Novices (n=10)	42.5% (28.5-57.9)	47.0% (26.2-68.9)	.07
P (experts vs novices)		.04	.08	

[a]D-VIA: digital visual inspection with acetic acid.

[b]D-VILI: digital visual inspection with Lugol's iodine.

[c]D-VIA vs D-VILI: vExperts vs. Clinical PDntion either "e either" re prefrred.

Figure 3. Sensitivity and specificity of digital visual inspection with acetic acid (D-VIA) and Lugol's iodine (D-VILI) for all raters (red triangle), experts (blue square), and novices (green circle).



Discussion

Principal Findings

Our study shows that the use of digital images for triaging HPV-positive women allows detection of most precancerous and cancerous lesions. When considering all raters, D-VIA showed a significantly higher sensitivity than D-VILI, whereas D-VILI was associated with significantly higher specificity. Furthermore, the specificity of D-VIA was higher among the more experienced screeners than among the less experienced ones. Although we did not compare these digital inspection methods with naked-eye techniques, this last finding could relate to the fact that bare-eye VILI is relatively easier to interpret than VIA, whose validity relies more on the experience and training level of the health worker [4].

As the overall sensitivity was higher for D-VIA, this study supports the superiority of D-VIA to D-VILI for CIN2+ detection. This result is partly in line with the WHO policy, according to which a primary visual approach should be VIA-based [3]. Furthermore, with a sensitivity of 94.1% and specificity of 50.4% for the detection of CIN2+, D-VIA showed a performance comparable with the one found in a study evaluating the performance of a low-cost magnifying device, the Magnivisualiser, where VIA had a sensitivity of 88.3% and a specificity of 55.8% [6].

Although VIA is currently recommended as a single screening tool, VILI is used to aid health workers in making the diagnosis because of its higher specificity and facility to interpret, especially for less-experienced health care providers. The cost and availability of Lugol's iodine, however, may potentially be an obstacle to the use of VILI in LMICs [7], resorting to the use of VIA alone in some regions.

In comparison to a recent meta-analysis on cervical cancer screening using naked-eye inspection methods which states that VILI seems to be the most sensitive test to use in the African continent, our study shows that D-VIA is more sensitive for detecting cervical lesions than D-VILI (94.1% vs 78.8%) [8]. The higher sensitivity obtained with D-VIA compared with D-VILI could be explained by the fact that, in this context, the raters were able to examine the images for a longer period of time and to compare the native photographs consecutively with those with D-VIA. This is not possible with the bare-eye approach because, once the application of Lugol's iodine is completed, the cervix appears as brown or black and the native and acetic acid appearance cannot be seen anymore. In addition, because the images were exclusively of HPV-positive women, the raters may have been inclined to give them a positive score in anticipation of a higher disease proportion in this group than in the general screening population.

In a low-income setting, such as that of Madagascar, cervical cancer screening tools that are commonly used in industrialized countries such as Pap smear and colposcopy, coupled with the continuous training of health care providers, are unlikely to be systematically available. The introduction of a smartphone-based approach for cervical cancer screening in such settings would allow to overcome some of the barriers to the implementation of screening programs in developing countries. The capture of cervical images with the smartphone camera allows the user to look back at either the native, post-VIA, or post-VILI cervix and to magnify the pictures in order to see them more closely before deciding whether or not to treat. In addition, the automatic saving of digital images on the smartphone allows the on-site, often less experienced and less qualified healthcare worker, to seek advice from long-distance off-site specialists [9].

Moreover, the use of automated phone applications is on the rise and might improve and facilitate CC screening in LMICs by providing a system to classify the images and to guide health workers through their decision-making algorithm [10]. Such mobile health tools are either free of charge or come at a very low price and can be easily installed on a smartphone without requiring any additional equipment. Their low cost and practicality distinguish them from other mobile colposcopy systems, such as the MobileODT (EVA system, enhanced visual assessment; Tel Aviv, Israel) for which the digital images' increased sharpness comes at the cost of a far more expensive and elaborate type of equipment. These aspects make the use of images taken by mobile phone a promising option for cervical cancer screening in low-income countries. Whether D-VIA and D-VILI should be used as a triage test for HPV-positive women or as a stand-alone screening tool is a question that yet remains unanswered. Further prospective studies are needed in order to assess the performance of D-VIA and D-VILI as a single, co-testing or triage screening tool.

Strengths and Limitations

The strengths of our study were the large number of consecutive cases included in a "triage context" and the fact that all of them underwent ECC and cervical biopsy, which served as the gold standard. In addition, patient recruitment and data collection took place in a real-world setting, which entails several consecutive, unselected cases.

Limitations that need to be addressed are the fact that D-VIA and D-VILI images of each patient were assessed separately rather than consecutively. Another limitation is the fact that most of our raters were medical doctors, which does not allow us to validate our results for all healthcare workers, such as nurses and midwives. The technical difficulties encountered in obtaining high quality images also did not allow us to include pictures of all the patients in the analysis. Additional practice is necessary in order to improve the quality of the images before transferring the smartphone use on to a real-life setting.

Conclusions

In conclusion, our study emphasizes the higher sensitivity of D-VIA when compared to D-VILI for detecting precancerous lesions, although at the cost of slightly lower specificity. Further research should focus on comparing the performance of these two digital screening techniques with naked-eye methods and on assessing their feasibility in a low-resource context.

Acknowledgments

The authors would like to thank the "Action Madagascar" Foundation and the University Hospitals of Geneva, Switzerland for their financial support of the study. The funders had no role in the study design, data collection, analysis, and in the preparation of the manuscript.

Conflicts of Interest

None declared.

References

1. Globocan. Globocan 2012 URL: http://globocan.iarc.fr/Default.aspx [accessed 2017-01-20] [WebCite Cache ID 6neUmeKm4]
2. McGraw SL, Ferrante JM. Update on prevention and screening of cervical cancer. World J Clin Oncol 2014 Oct 10;5(4):744-752 [FREE Full text] [doi: 10.5306/wjco.v5.i4.744] [Medline: 25302174]

3. World Health Organization. WHO Guidelines for screening and treatment of precancerous lesions for cervical cancer prevention. Geneva, Switzerland: WHO; 2013.
4. Fokom-Domgue J, Combescure C, Fokom-Defo V, Tebeu PM, Vassilakos P, Kengne AP, et al. Performance of alternative strategies for primary cervical cancer screening in sub-Saharan Africa: systematic review and meta-analysis of diagnostic test accuracy studies. BMJ 2015 Jul 03;351:h3084 [FREE Full text] [Medline: 26142020]
5. Ricard-Gauthier D, Wisniak A, Catarino R, Van Rossum AF, Meyer-Hamme U, Negulescu R, et al. Use of smartphones as adjuvant tools for cervical cancer screening in low-resource settings. J Low Genit Tract Dis 2015 Oct;19(4):295-300. [doi: 10.1097/LGT.0000000000000136] [Medline: 26247260]
6. Singh V, Parashari A, Gupta S, Sodhani P, Sehgal A. Performance of a low cost magnifying device, Magnivisualizer, versus colposcope for detection of pre-cancer and cancerous lesions of uterine cervix. J Gynecol Oncol 2014 Oct;25(4):282-286 [FREE Full text] [doi: 10.3802/jgo.2014.25.4.282] [Medline: 25142620]
7. Horo AG, Didi-Kouko Coulibaly J, Koffi A, Tchounga B, Seni K, Aka KE, et al. Cervical cancer screening program by visual inspection: acceptability and feasibility in health insurance companies. Obstet Gynecol Int 2015;2015:798453 [FREE Full text] [doi: 10.1155/2015/798453] [Medline: 26167178]
8. Qiao L, Li B, Long M, Wang X, Wang A, Zhang G. Accuracy of visual inspection with acetic acid and with Lugol's iodine for cervical cancer screening: meta-analysis. J Obstet Gynaecol Res 2015 Sep;41(9):1313-1325. [doi: 10.1111/jog.12732] [Medline: 26014371]
9. Catarino R, Vassilakos P, Scaringella S, Undurraga-Malinverno M, Meyer-Hamme U, Ricard-Gauthier D, et al. Smartphone use for cervical cancer screening in low-resource countries: a pilot study conducted in Madagascar. PLoS ONE 2015 Jul 29;10(7):e0134309. [doi: 10.1371/journal.pone.0134309]
10. Kassianos AP, Emery JD, Murchie P, Walter FM. Smartphone applications for melanoma detection by community, patient and generalist clinician users: a review. Br J Dermatol 2015 Jun;172(6):1507-1518. [doi: 10.1111/bjd.13665] [Medline: 25600815]

Abbreviations

CIN1: cervical intraepithelial neoplasia grade 1
CIN2: cervical intraepithelial neoplasia grade 2
CIN2+: cervical intraepithelial neoplasia grade 2 or worse
CIN3: cervical intraepithelial neoplasia grade 3
D-VIA: digital smartphone-based visual inspection of the cervix with acetic acid
D-VILI: digital smartphone-based visual inspection of the cervix with Lugol's iodine
ECC: endocervical curettage
HPV: human papillomavirus
LMICs: low- and medium-income countries
VIA: visual inspection of the cervix with acetic acid
VILI: visual inspection of the cervix with Lugol's iodine
WHO: world health organization

Edited by G Eysenbach; submitted 28.09.16; peer-reviewed by M Schiffman, KE Yeates, EAW Shemer, A Goodman; comments to author 29.12.16; revised version received 30.01.17; accepted 10.04.17; published 29.05.17.

Please cite as:
Urner E, Delavy M, Catarino R, Viviano M, Meyer-Hamme U, Benski AC, Jinoro J, Heriniainasolo JL, Undurraga M, De Vuyst H, Combescure C, Vassilakos P, Petignat P
A Smartphone-Based Approach for Triage of Human Papillomavirus-Positive Sub-Saharan African Women: A Prospective Study
JMIR Mhealth Uhealth 2017;5(5):e72
URL: http://mhealth.jmir.org/2017/5/e72/
doi:10.2196/mhealth.6697
PMID:28554879

Original Paper

Text Message Feedback to Support Mindfulness Practice in People With Depressive Symptoms: A Pilot Randomized Controlled Trial

Susanne Kraft[1], Dipl Psych; Markus Wolf[2], Dipl Psych, PhD; Thomas Klein[1], MSc (Psych); Thomas Becker[1], MD; Stephanie Bauer[3], Dipl Psych, PhD; Bernd Puschner[1], Dipl Psych, PhD

[1]Section Process-Outcome Research, Department of Psychiatry II, Ulm University, Günzburg, Germany

[2]Clinical Psychology and Psychotherapy Research, Institute of Psychology, University of Zurich, Zurich, Switzerland

[3]Center for Psychotherapy Research, University Hospital Heidelberg, Heidelberg University, Heidelberg, Germany

Corresponding Author:
Bernd Puschner, Dipl Psych, PhD
Section Process-Outcome Research
Department of Psychiatry II
Ulm University
3rd floor
Ludwig-Heilmeyer-Str 2
Günzburg, 89312
Germany
Phone: 49 8221 982866
Fax: 49 8221 9628160
Email: bernd.puschner@bkh-guenzburg.de

Abstract

Background: It has been shown that mindfulness practice can be helpful in preventing relapse from depression. However, practicing mindfulness regularly at home is often a challenge for people with depression. Mobile phone text messaging (short message service, SMS) may be a feasible approach to assist regular mindfulness home practice.

Objective: The aim of this study was to evaluate the feasibility of text message–based feedback to support mindfulness practice in people with depressive symptoms after inpatient psychiatric treatment.

Methods: Participants received a manualized group introduction to three mindfulness exercises during inpatient treatment and were randomized at hospital discharge. All participants were asked to practice the exercises daily during the 4-month follow-up period. Only participants allocated to the intervention group received reinforcing feedback via mobile phone text messages after reporting their mindfulness practice via text message. Participation rates and satisfaction with the interventions were evaluated, and effects on relevant outcomes were explored.

Results: Of the 176 eligible inpatients invited to participate, 65.9% (116/176) attended the introductory mindfulness group at least once, 33.0% (58/176) were willing to participate in the study, and 41 were randomized. The majority 85% (35/41) of these participants completed the study. Among the participants allocated to the intervention group (n=21), 81% (17/21) used the text message support at least once. The average number of text messages sent during the intervention period was 14 (SD 21, range 0-91). Satisfaction rates were high. Preliminary analyses of the effects of the intervention yielded mixed results.

Conclusions: Findings indicate that text messaging following inpatient treatment is feasible for some, but not for all people with depressive symptoms. Modest use of the text messaging intervention and its mixed effects imply that dose and ingredients of the intervention should be increased for this group of patients in a future full-size RCT. Such a larger study should also include a process evaluation to investigate moderators of the effect of mindfulness practice and text message feedback on clinical outcome.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 58808893; http://www.controlled-trials.com/ISRCTN58808893 (Archived by Webcite at http://www.webcitation.org/6pmrDRnGt)

(JMIR Mhealth Uhealth 2017;5(5):e59) doi:10.2196/mhealth.7095

KEYWORDS

mindfulness; text messaging; pilot study; randomized controlled trial

Introduction

A recent meta-analysis showed that mindfulness-based interventions contribute to reducing depressive symptoms (*d*=0.30 at 2 months; *d*=0.23 at 3-6 months follow-up [1]) and may be beneficial in the treatment of outpatients with acute depression [2]. Although some studies evaluated mindfulness-based interventions for inpatients with psychosis or borderline personality disorder (eg, [3,4]), to our knowledge, only one study has evaluated mindfulness techniques in depressed inpatients [5]. In this uncontrolled pilot study, an 8-session mindfulness program yielded significant pre-post changes in depression and mindfulness. However, attrition was considerable with only half of the participants completing the intervention.

The continuous practice of mindfulness exercises is a key component of mindfulness-based interventions. Time spent on formal mindfulness practices, such as body scan and mindfulness on breathing, is positively associated with outcome [6]. However, with a large share (38%) of people with depression who received mindfulness-based cognitive therapy (MBCT) not regularly practicing at home, lack of homework compliance is a frequent problem [7].

Mobile health (mHealth) apps might increase the effectiveness of mindfulness-based interventions by improving homework compliance. Being simple and efficient, the mobile phone short message service (SMS) is increasingly used to assist the delivery of mental health care [8]. A growing number of studies show that various apps of texting (monitoring, feedback, communication, homework reminders) contribute to improving uptake and outcome of mental health care [9,10].

Pilot studies have evaluated the use of texting to support homework assignments for people with depression receiving cognitive behavioral therapy [11], and as part of a Web-based monitoring and feedback intervention with a focus on mindfulness and acceptance [12]. Moreover, although recent studies have evaluated the effectiveness of Web-based and mobile phone apps, some of them also with a focus on mindfulness [13-18], there is a lack of research on the feasibility of texting interventions to support mindfulness practice in people with depression.

Thus, we developed a low-intensity program using texting to support postdischarge mindfulness practice in people with depressive symptoms receiving inpatient treatment. This randomized controlled pilot study investigated (1) participation rates during the different stages of the study and the intervention (recruitment, introductory group, randomization, texting, mindfulness exercises during follow-up, return rates of follow-up questionnaires); (2) satisfaction with the mindfulness training and the text message intervention; and (3) feasibility of the outcome measures.

Methods

Design

Recruitment for the study "An SMS-Assisted Mindfulness-based Intervention for Relapse Prevention in Depression" (MIND-S, ISRCTN58808893) took place between September 2013 and June 2014 at Ulm University's Department of Psychiatry II in Günzburg, Germany. The hospital provides acute and long-term inpatient mental health care for a catchment area of about 671,000 inhabitants in rural Bavaria.

MIND-S was a pilot two-arm randomized clinical trial. Participants were invited to attend mindfulness group sessions at the hospital before discharge. Randomization took place shortly before discharge, with participants allocated to the intervention group receiving a simple texting intervention to support mindfulness exercises at home. Data was collected at three measurement points: (1) baseline (after giving informed consent, before or shortly after the first mindfulness group session), (2) prerandomization (shortly before discharge), and (3) follow-up (4 months after discharge). The study was approved by Ulm University's Ethics Committee.

Participants

Participants were included if they were inpatients or day patients aged 18-75 years, and showed symptoms of depression according to the clinical judgment of their inpatient therapist. Exclusion criteria were the presence of psychotic symptoms or a history of schizophrenia, current manic state, risk of a dissociative crisis, severe cognitive impairment, persistent severe substance abuse, suicidality or risk of self-harm during the current illness episode, insufficient command of the German language, and lack of a mobile phone. To identify eligible patients, therapists were informed about the study and the hospital database was regularly monitored. Therapists of possibly eligible patients were asked to give a more detailed account regarding inclusion and exclusion criteria. Eligible patients were contacted personally by the first author SK and provided with oral and written information about the study, asked to give informed consent and to complete the baseline questionnaire, and invited to attend the introductory mindfulness group. As the mindfulness group was open to all inpatients, several participants asked to visit the group once, before they gave their informed consent, and completed the baseline questionnaire shortly (maximum 1 day) afterwards.

Randomization

After participants returned the prerandomization questionnaire, they were randomly allocated at a 1:1 ratio to the intervention or the control group. Randomization was based on a centralized procedure, which was coordinated by the Heidelberg site independent of the recruiting clinical site (Günzburg) using computer-generated random numbers.

Intervention

Part 1: Mindfulness Group Training During Inpatient Treatment

The mindfulness instructions used in this study were based on the exercises suggested by Segal and colleagues [19] on mindful breathing, mindful walking, and the "body scan." This program ("MBCT") was originally developed for formerly depressed outpatients. Since participants were acutely depressed, all exercises were shortened to a maximum duration of 10 min and guided throughout. Following a general introduction on the principles of mindfulness and self-compassion in the context of depression, each of the following guided exercises was practiced in the group for 5-10 min: (1) mindful breathing, (2) mindful walking, and (3) mindfulness of the body ("body scan," with a focus on feet and legs). Each participant received a short written description of the exercises to take home after discharge from hospital, and it was recommended to practice one or more of the exercises about once a day for at least 5 min.

The mindfulness group training was provided by the first author, who is a licensed cognitive behavioral psychotherapist with 7 years of experience in both inpatient group therapy and mindfulness practice, and who completed introductory courses in mindfulness-based treatments (but did not receive a full training in mindfulness-based stress reduction or MBCT). The 60-min sessions took place weekly, with the manualized contents described above being repeated every time. New participants could join the group at any time. The group was offered as a part of the standard clinical treatment and was open also to patients who were not participating in the study. However, therapists were asked to only send patients who match the study criteria. Study participants were required to attend at least once in order to continue the study.

Part 2: Texting After Hospital Discharge

Participants allocated to the intervention group were asked to send a text message via their mobile phones to the study center whenever they practiced one or more of the mindfulness exercises. The text message should contain information on kind (A for "Atmen," ie, mindful breathing; G for "Gehen," ie, mindful walking; K for "Körper;" ie, mindfulness of the body) and duration of the exercise (minutes exercised), resulting in a short alphanumeric code (eg, A10 for a 10-min practice of mindful breathing, or G5K5 for 5 min of mindful walking followed by a 5-min body scan). As a reminder, each participant received a brief pocket guide explaining the code translations to be sent via text message. After sending a text message, the participant received an automated reply, which consisted of (1) a brief positive reinforcing feedback that was drawn from a pool of 86 messages (eg, "Great! Try to be kind to yourself while practicing."), which were formulated in advance by the study team based on the MBCT literature and randomly assigned by a computer program; and (2) the total time (in minutes) the participant had practiced since the beginning of the text message intervention (eg, "You have already practiced 25 min so far"). If a participant's text message did not match the required format, he or she received a message on how to use the program. A reminder was sent if no text messages were received for more than 1 week. Reminders were continuously sent every week until the end of the intervention in case of persistent nonresponse. We decided to send reminders weekly instead of daily because the main function of the messages was to gently reinforce training behavior (self-management) rather than merely reminding patients to practice the mindfulness exercises. SK introduced participants to the text message intervention, which started immediately after discharge from hospital and lasted 4 months. The few participants with no texting experience received instructions on how to send and receive text messages. The intervention manual is available from the authors upon request.

Participants of the control group attended the mindfulness training group and were asked to regularly practice the exercises at home. They did not receive text message assistance during follow-up. The intervention was an add-on to treatment as usual. There were no constraints for participants to utilize any other treatment during the study period. All participants received €25 after returning the follow-up questionnaire. Additionally, participants in the intervention group received €10 at hospital discharge to cover their costs for sending text messages.

Measures

Outcomes of the intervention to be expected include broader effects on depression (severity of depressive symptoms and perseverative thinking) and on aspects more closely related to the nature of the intervention (mindfulness and self-compassion).

Severity of depressive symptoms was measured with the German version of the Brief Patient Health Questionnaire-9 Item (PHQ-9 [20]), which assesses the DSM-IV criteria for major depression via patient self-report with 9 items on a 4-point Likert scale ranging from 0 ("not at all") to 3 ("nearly every day"). The sum score ranges from 0 to 27 with higher scores indicating more severe depression.

The German version of the Perseverative Thinking Questionnaire (PTQ [21]) assesses *repetitive negative thinking* without referring to depressive symptoms in the item formulation. This self-report questionnaire consists of 15 items answered on a 5-point Likert scale ranging from 0 ("never") to 4 ("almost always"), which describe how participants typically think about negative experiences or problems (eg, "My thoughts repeat themselves"). The total score ranges from 0 to 60, with higher scores indicating more repetitive thinking.

Mindfulness was assessed with the 14-item short form of the Freiburg Mindfulness Inventory (FMI; German: "Freiburger Fragebogen zur Achtsamkeit," FFA-14 [22]). Items are rated on a 4-point Likert scale ranging from 1 ("rarely") to 4 ("almost always"). The sum score ranging from 14 to 56 was used, with higher scores indicating more mindfulness.

Self-compassion was assessed with the German version (SCS-D [23]) of the short form of the Self-Compassion Scale (SCS-SF [24]). The 12 items (eg, "I'm disapproving and judgmental about my own flaws and inadequacies") are rated on 5-point Likert scale ranging from 1 ("very rarely") to 5 ("very often"). Higher mean total scores indicate more self-compassion.

A number of instruments measured feasibility and acceptance of the intervention at the various stages of the trial. First,

mindfulness practice was measured via a short self-constructed questionnaire at follow-up. Participants were asked how often they practiced each of the three mindfulness exercises during each of the 4 months of the intervention period, yielding total number of exercises. Also at follow-up, participants were asked how long on average they practiced each of the three exercises in each of the 4 months. The mean duration per exercise was multiplied with the number of times this exercise was practiced in a given month, and this information was summarized for the three exercises over 4 months, yielding total duration of home practice per exercise. If participants reported an exercise without specifying the time spent on exercising, missing data was replaced by the grand means based on available data for each type of exercise (9 min for mindfulness breathing, 11 min for mindfulness walking, and 14 min for mindfulness of the body). Extreme outliers were omitted (3 participants in the control group who reported to have practiced over 200 times). Second, at baseline , *willingness to send text messages* ("Do you think you would send a text message after practicing?") and *previous experience with the texting and mindfulness exercises* were assessed with 1-item questions each. Third, *satisfaction with the text message–based intervention* was assessed at follow-up with a 20-item adaptation of a questionnaire developed and used in earlier research of our group [25], asking on a 5-point rating scale ("I do not agree," "I rather not agree," "I somewhat agree," "I fairly agree," and "I totally agree") for general satisfaction, satisfaction with the text message feedback, and technical problems. Fourth, *satisfaction with the mindfulness intervention* was assessed via a self-constructed 10-item questionnaire at discharge and follow-up, asking on the same 5-point rating scale about satisfaction with the mindfulness introduction, as well as about potential problems with and subjective effects of the mindfulness exercises. Fifth, at follow-up, participants of both groups were asked to rate the *usefulness* of the intervention as a whole ("yes, it helped me a lot," "yes, it helped me somewhat," "it neither helped nor harmed me," "no, it rather harmed me," and "no, it harmed me a lot"), including the option to add open-ended comments and suggestions.

Sociodemographic information was assessed at baseline. Furthermore, inpatient therapists were asked to document the *diagnoses* of their participating patients at study intake according to ICD-10.

Data Analysis

According to recommendations for reporting results of pilot studies [26,27], reporting of results focuses on descriptive statistics, that is, mean and standard deviation for continuous variables, and frequencies and percentages for categorical variables. Feasibility indicators include numbers of patients eligible, willing to participate in the study, and to be randomized; numbers of participants lost to follow-up; intended, and actual use of the text message intervention; the number of text messages sent during the intervention; as well as frequency and duration of home practice.

Chi-square and *t*-tests were used to analyze differences of feasibility indicators by allocation. For the total scores of PHQ-9, PTQ, FMI, and SCS-D group by time interaction effects were tested using repeated measures analysis of variance. Additionally, effect sizes *d* (standardized mean between-group differences corrected for prerandomization differences) were calculated for all outcome measures.

Results

Sample

Socioeconomic status was assessed at baseline (see Table 1). Participants were 44-years-old on average, most of them were female, married, or living together with a partner (see Table 1). Level of education was predominantly low, and most were working at least part-time. Mean duration since the first occurrence of depressive symptoms was 11 years. According to their PHQ-9 score at baseline, more than three quarters (81.1%) of the patients showed substantial symptoms of depression (PHQ-9 ≥10). At baseline, participants who were later allocated to the intervention group were less depressed and showed more self-compassion than participants of the control group. The majority of the participants had at least some experience with texting. Regarding mindfulness exercises, about half had tried mindfulness practice once, a few were practicing regularly for up to 5 years.

Participation in the Introductory Group and the Study

It was found that 54.8% (176/321) of the screened patients were judged as eligible and invited to participate in the study. Of these, 65.9% (116/176) visited the introductory group at least once. The mean number of group visits of all participants (later randomized or not) was 2.4 (SD 1.8, range 1-9), the number of group attendees varied between 1 and 19 (mean 6.9, SD 3.5). In total, 33.0% (58/176) of the invited patients consented to participate in the study and completed baseline measures. Of the latter, 71% (41/58) completed prerandomization measures and were randomized. Randomized participants took part in the introductory group 3.4 times (SD 2.0) on average (intervention: mean 3.3, SD 2.0; control: mean 3.5, SD 2.1). Follow-up data were available for 85% (35/41) of the randomized participants, with no differences by allocation. Figure 1 depicts the (simplified) flow of participants through the stages of the trial.

Most common reasons to decline study participation (not systematically assessed) were not wanting to complete questionnaires (5%; 3/57) or using the mobile phone (4%; 2/57), a lack of interest in mindfulness practice (5%; 3/57), and feeling that participation in the study would be "too much" (5%; 3/57). Most patients (70%; 40/57) did not give any reason. Still, of the 32.4% (57/176) of patients who were not interested in study participation, 54% (31/57) attended the mindfulness training group at least once.

Table 1. Characteristics of study participants.

Characteristics	IG[a] (n=21)	CG[b] (n=20)	Differences
Age (in years), mean (SD[c])	43.4 (12.7)	44.5 (13.5)	$t_{39}=-0.26$; $P=.80$
Gender (female), n (%)	13 (61.9)	15 (75.0)	$\chi^2_1=0.8$; $P=.37$
Level of education			
Qualification for university entrance, n (%)	3 (14.3)	4 (20.0)	$\chi^2_1=0.2$; $P=.63$
Lower qualification, n (%)	18 (85.7)	16 (80.0)	
Marital status			
Married or living together with partner, n (%)	10 (47.6)	11 (45.0)	$\chi^2_2=0.4$; $P=.84$
Single, n (%)	7 (33.3)	5 (25.0)	
Other, n (%)	4 (19.0)	4 (20.0)	
Employment status			
Full-time, n (%)	8 (38.1)	6 (30.0)	$\chi^2_3=0.5$; $P=.93$
Part-time, n (%)	5 (23.8)	5 (25.0)	
Unemployed, n (%)	6 (28.6)	6 (30.0)	
Other, n (%)	2 (9.5)	3 (15.0)	
Primary diagnosis			
Major depression[d], n (%)	16 (76.2)	19 (95.0)	$\chi^2_1=2.9$; $P=.09$
Other[e], n (%)	5 (23.8)	1 (5.0)	
Illness duration (years), mean (SD)	10.8 (12.2)	11.7 (8.4)	$t_{38}=-0.29$; $P=.78$
PHQ-9[f] (depressive symptoms), mean (SD)	12.74 (5.69)	18.61 (4.86)	$t_{35}=-3.37$; $P=.002$
PTQ[g] (perseverative thinking), mean (SD)	38.50 (10.97)	45.32 (9.95)	$t_{35}=-1.98$; $P=.06$
FMI[h] (mindfulness), mean (SD)	30.56 (5.53)	27.16 (4.91)	$t_{35}=1.98$; $P=.06$
SCS-D[i] (self-compassion), mean (SD)	2.53 (0.76)	2.05 (0.59)	$t_{36}=2.15$; $P=.039$
Experience with mindfulness exercises			
None, n (%)	7 (35.0)	10 (50.0)	$\chi^2_2=1.7$; $P=.44$
Tried once, n (%)	12 (60.0)	8 (40.0)	
Experienced, n (%)	1 (5.0)	2 (10.0)	
Experience with texting			
None n (%)	5 (23.8)	3 (15.0)	$\chi^2_2=3.2$; $P=.20$
Some experience, n (%)	4 (19.0)	9 (45.0)	
Very experienced, n (%)	12 (57.1)	8 (40.0)	

[a]IG: intervention group.

[b]CG: control group.

[c]SD: standard deviation.

[d]ICD-10=F32.1, F32.2, F33.1, or F33.2.

[e]ICD-10=F31.4, F40.01, F41.0, F43.2, or F61.0.

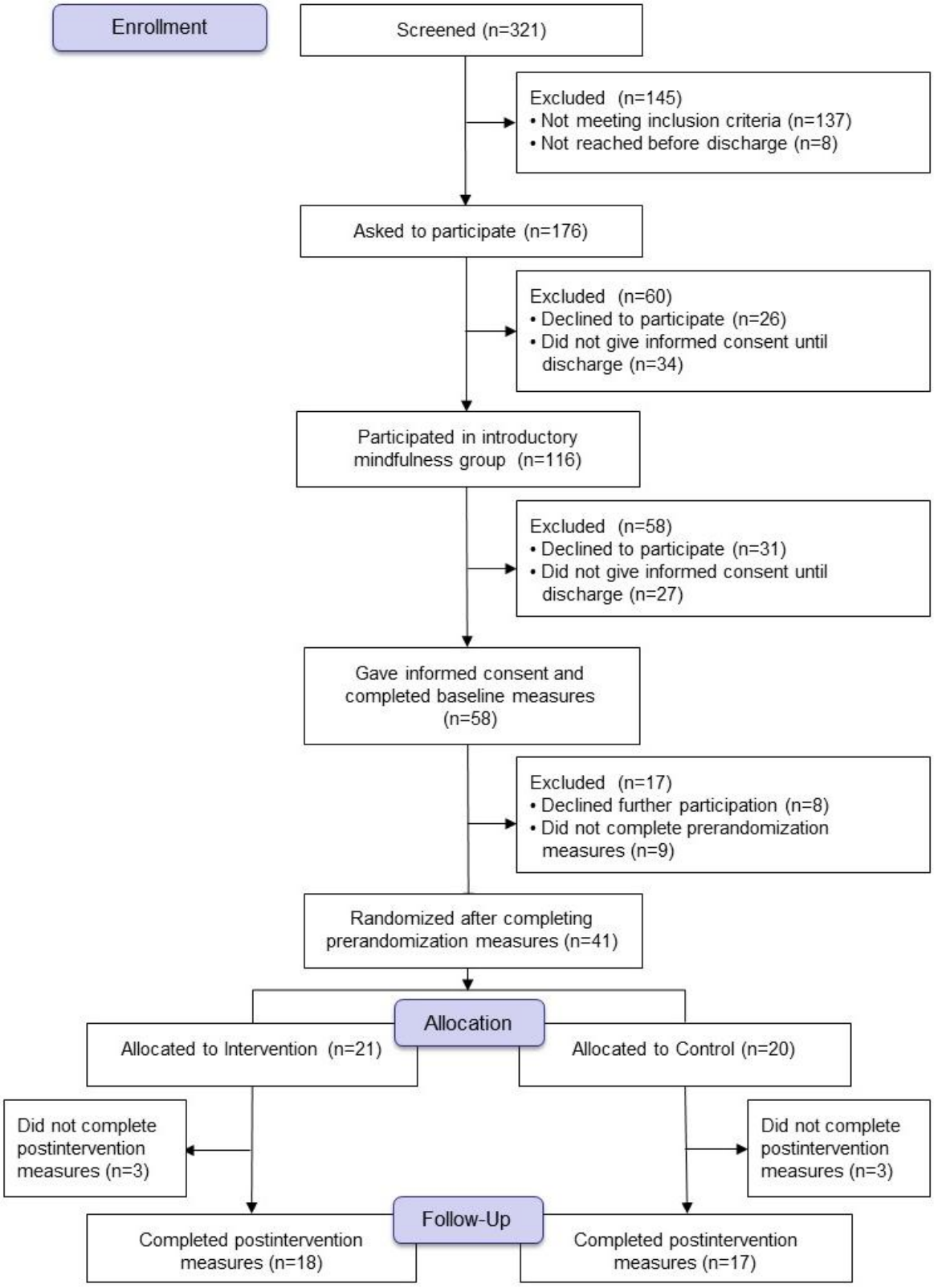
[f]PHQ-9: Patient Health Questionnaire.

[g]PTQ: Perseverative Thinking Questionnaire.

[h]FMI: Freiburg Mindfulness Inventory.

[i]SCS-D: Self-Compassion Scale; Missing values: illness duration: N=1 (IG), PHQ-9: N=2 (each IG and CG); PTQ, FMI: N=3 (IG), N=1 (CG); SCS-D: N=2 (IG), N=1 (CG).

Figure 1. Participant flow through the stages of the trial.



Participation in the Text Message Feedback Intervention

Before randomization, 59% (24/41) of the study participants expressed their intent to send text messages after practicing a mindfulness exercise if randomized into the intervention group. It was found that 20% (8/41) of the participants indicated that they would not and 22% (9/41) stated that they would "rather not" use the text message assistance. Of the 24 participants expressing intent to send text messages, 50% (12/24) were later

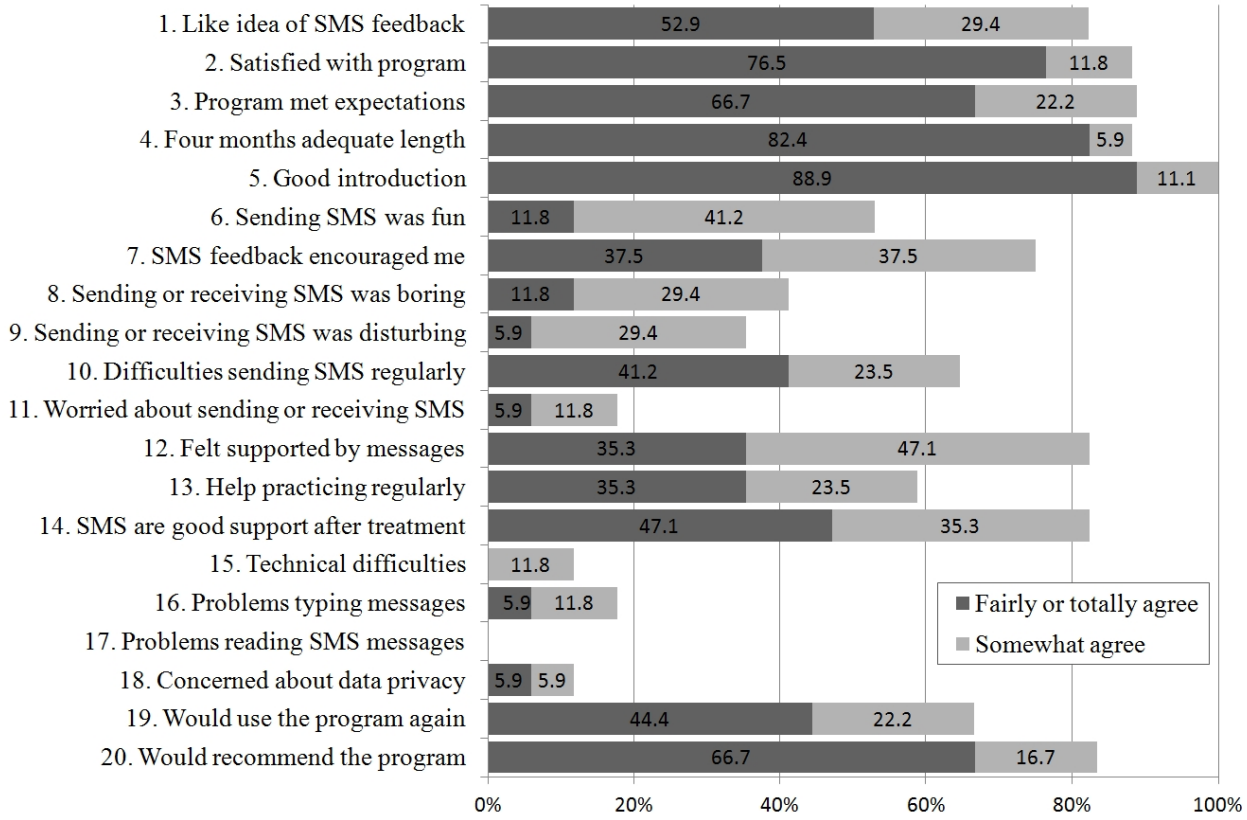
assigned to the intervention group. At follow-up, about two-third of the participants in the intervention group reported that they had not (18%) or not always (47%) sent a text message after mindfulness practice, whereas about one-third reported that they had done so always (12%) or most of the time (24%).

During the intervention period, participants sent 294 text messages reporting 395 exercises. Of the 21 participants, 81% (17/21) in the intervention group sent at least one text message and 67% (14/21) of participants texted more than once. Two participants sent at least one message a week. On average, participants sent 14.00 (SD 21.00, median 9, range 0 - 91) messages, reporting 18.81 (SD 5.44, median 12) exercises during 4 months, that is, about one exercise per week. Correlations of the number of text messages with several parameters are reported elsewhere [28]. The mean duration per exercise reported via text message was 9.94 min (SD 5.38) with mindfulness of the body exercises showing the longest duration (mean 13.84, SD 4.13, N=103), followed by mindful breathing (mean 8.95, SD 5.48, N=211), and mindful walking (mean 7.54, SD 3.70, N=81).

Satisfaction with the Interventions

Findings about satisfaction with the text message intervention are shown in Figure 2. Overall, participants showed high satisfaction. Specifically, the program met the expectations of most of the participants, more than half reported that it helped them to practice the mindfulness exercises regularly, and most felt generally supported by the text message feedback. Two-third reported that they would use the program again, and the majority indicated that they would recommend it to a friend. However, about two-third of the participants stated that they found it difficult to send messages on a regular basis, and some were concerned about data privacy.

Figure 2. Satisfaction with the text message intervention.



Notes: SMS=short message service. Missing values: items #1, #2, #4, #6, #8-18: N=1; #7: N=2. Participants who did not or rather not agree not shown.

At follow-up, participants of both groups were asked to evaluate the mindfulness exercises. The introduction into the mindfulness exercises at the hospital appeared appropriate and clear to all (see Figure 3). While the majority of the intervention group found that they were able to practice mindfulness in everyday life, this was the case for less than half of the participants of the control group. Furthermore, in comparison to the control group, more participants in the intervention group perceived the mindfulness exercises helpful in relationships and in coping with rumination and negative feelings. Participants in the control group reported more difficulties in practicing the mindfulness exercises alone.

At follow-up, 67% (35/53) of the participants in both intervention and control groups reported that taking part in the study helped them to some degree, the rest said it neither helped nor harmed them. The rate of agreement was higher in the intervention group (78%) than in the control group (53%). Of the 11 participants who provided written feedback at follow-up, 2 stated that they often forgot to send a text message, another 2 wrote that the weekly text message reminders were perceived as helpful, and 1 recommended further reminders in larger

intervals. One participant stated that practicing in the group was easier than alone, and another one suggested to provide audio recordings of the exercises, and finally one participant recommended to offer the inpatient group sessions more frequently.

Figure 3. Satisfaction with the mindfulness introduction and exercises at follow-up.

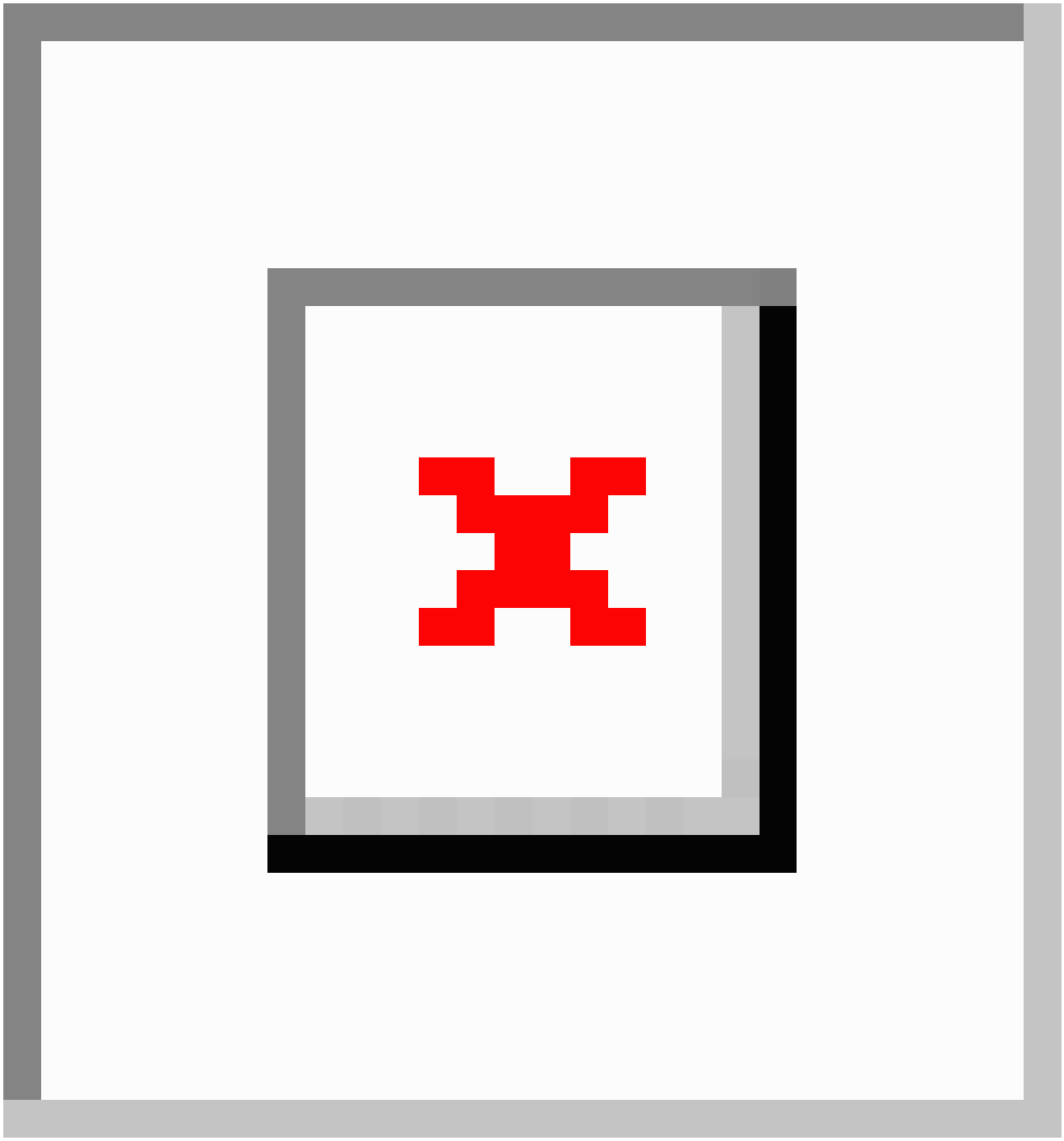


Table 2. Effect of the intervention on mindfulness practice and outcomes (N=35).

Outcome measures	Prerandomization		4-month follow-up		Test statistics	*P* value	Effect size (95% CI)
	IG[a], mean (SD[b])	CG[c], mean (SD)	IG, mean (SD)	CG, mean (SD)			
Number of exercises practiced	7.88 (4.92)	10.06 (6.80)	62.67 (61.64)	48.71 (44.28)	t_{30}=−0.72	.48	0.25 (−0.45 to 0.96)
Total time practiced (min)	23.75 (13.11)	35.13 (16.55)	642.17 (815.26)	579.50 (784.28)	t_{30}=−0.22	.83	0.08 (−0.62 to 0.78)
PHQ-9[d] (depressive symptoms)	37.47 (6.49)	32.00 (6.83)	8.94 (6.61)	12.06 (7.24)	$F_{1,31}$=0.17	.68	0.14 (−0.53 to 0.82)
PTQ[e] (perseverative thinking)	3.10 (0.77)	2.61 (0.85)	26.25 (19.28)	33.19 (16.33)	$F_{1,30}$=0.85	.37	−0.26 (−0.84 to 0.31)
FMI[f] (mindfulness)			36.40 (8.72)	32.53 (6.40)	$F_{1,28}$=0.44	.51	0.25 (−0.49 to 0.99)
SCS-D[g] (self-compassion)			3.11 (1.06)	2.63 (0.71)	$F_{1,29}$=0.01	.94	0.02 (−0.59 to 0.63)

[a]IG: intervention group.

[b]SD: standard deviation.

[c]CG: control group.

[d]PHQ-9: Patient Health Questionnaire-9 Item.

[e]PTQ: Perseverative Thinking Questionnaire.

[f]FMI: Freiburg Mindfulness Inventory.

[g]SCS-D: Self-Compassion Scale; Missing values: PHQ-9: N=1 (each IG and CG); PTQ: N=2 (IG), N=1 (CG); FMI: N=3 (IG), N=2 (CG); SCS-D: N=2 (each IG and CG).

Feasibility of the Outcome Measures

The effects of the post-hoc assessment of mindfulness home practice (number and duration of exercises practiced) and the four selected outcome questionnaires (PHQ-9, PTQ, FFA, and SCS-D) yielded near zero to small, nonsignificant effect sizes (see Table 2). Note that, due to different scoring, a positive effect size on the symptom measures (PTQ, PHQ) indicates an outcome in favor of the intervention group, whereas on the FMI and the SCS-D, a positive effect size indicates a result in favor of the control group.

The largest but still small differences were found on perseverative thinking, self-reported mindfulness, and the number of reported exercises. Effects for the former two measures were against expectations because the control group scored lower on perseverative thinking (d=−0.26) and higher on mindfulness (d=0.25) than the intervention group. The number of exercises practiced tended to be higher in the intervention group (d=0.25), and there was small effect on the PHQ-9 in favor of the intervention group.

Discussion

Principal Findings

To our knowledge, this is the first study to evaluate the feasibility of a simple, low-intensity texting app to support mindfulness practice in people with depressive symptoms after inpatient psychiatric treatment. The study showed that a considerable proportion of people with depressive symptoms receiving in inpatient treatment are interested in a mHealth mindfulness program, and that most participants were satisfied and experienced the intervention as helpful.

Participation in the Introductory Group and the Study

According to introductory group participation rates, the mindfulness exercises seemed to interest about two-third of the target population of psychiatric inpatients with depressive symptoms. Only half of them could be motivated to participate in the study (eg, fill out questionnaires and participate in the texting intervention). However, once they decided to participate in the study, most patients stayed in the trial until the end.

Retention rates were 71% for the first (prerandomization) and 85% for the second (postrandomization) part of the study, falling within the upper range of other mHealth studies (43-100% [29]). Similar, sometimes higher, attrition rates were found in studies evaluating mindfulness-based interventions in depressed patients delivered face-to-face (49% [5]; 8-38% [2]) or via the Internet or mobile phone (57% [14]; 38% [15]). Good retention rates speak to the feasibility of the study design and committed study staff. However, the possibility of a selection bias toward including rather motivated and compliant patients cannot be ruled out.

Participation in the Text Message Feedback Intervention

On average, participants sent about one text message per week which falls behind expectations, as daily practice was recommended. However, this result is not surprising, as before the beginning of the intervention, less than 60% of the randomized participants indicated that they intended to text. Comparing the postintervention self-report with the information from the text messages, participants texted only about every third time after a mindfulness exercise. Likewise, about two-third of the participants reported that they had not or not always sent a text message after they had practiced, and that

they had difficulties texting regularly. However, the possibility of a combined effect of a social desirability bias and a 4-month recall bias should be taken into account. Nevertheless, participation patterns are comparable with another study which examined text-messaging support in the treatment of people with depression (65% response rate to text message reminders [11]), indicating that expectations might have been unrealistic. Another explanation could be that a considerable number of participants did not feel the need to be supported every time they practiced. Taken together, these findings suggest a need to enhance mode of delivery and content of text messages including tailoring feedback to increase subjective meaning. This might also be achieved by sending daily queries inquiring about type and duration of practice.

Satisfaction With the Interventions

Participants of the intervention group were predominantly satisfied with the texting intervention. Most felt supported and encouraged by the messages, and over 80% would further recommend the program. Although use of the texting intervention was moderate, in absolute values, considerably more patients in the intervention than in the control group perceived the mindfulness exercises as helpful, especially regarding coping with rumination and negative feelings. More patients in the intervention group than in the control group stated that they had practiced in daily life. Furthermore, compared with the control group, in the intervention group, fewer patients at least fairly agreed that they had problems practicing alone. Although these are no significant results and should be regarded with caution, they might indicate that the text message feedback supports mindfulness home practice and could increase the effects of mindfulness exercises as intended [30].

Feasibility of the Outcome Measures

Although the main outcome measures (amount of home practice, depressive symptomatology) depict a small effect of the intervention, the additional measures (rumination, mindfulness, self-compassion) yielded zero or small effects in favor of the control group. Due to the small sample size, no final conclusion can be drawn. However, these results indicate that either the questionnaires were insensitive to changes, the dosage of intervention was not sufficient, or the observation period was too short to detect changes. It could be hypothesized that mindfulness and self-compassion are rather trait-like, less susceptible to change, and thus, need a longer and more intense intervention to change.

Limitations

This study has a number of limitations. First, the introduction to the mindfulness exercises and the text message intervention was provided by the first author, who also was involved in data collection and analysis. Second, the intervention was only implemented at one hospital, suggesting limited generalizability of findings. Third, there could have been a selection bias, as the inpatient therapists could recommend a patient to participate or not to participate in the study due to their subjective judgment. Fourth, the mindfulness introduction was a shortened and adopted version of other mindfulness-based programs, and has not been validated in this population. Fifth, there might be a recall bias regarding the post hoc assessment of the mindfulness exercises practiced in the follow-up period, combined with a social desirability bias. There might have been a tendency to over-report the mindfulness practice in the intervention group. Finally, due to the small sample size, all results of this pilot study should be interpreted with caution.

Conclusions

Taken together, the positive evaluations of the MIND-S program by the participants indicate that, in general, mindfulness practice augmented by text message feedback is a feasible intervention. However, the moderate use of the texting intervention and the mixed effects imply that dose and ingredients of the intervention should be increased for this group of patients in a future, full-scale RCT. Additional components might include expert or peer support such as regular mindfulness group visits after discharge, audio or video material, more frequent reminders, or individual expert support via texting, email or chat-groups. To minimize recall bias, frequency and duration of mindfulness practice should be assessed more frequently during the follow-up period. Furthermore, the intervention should also be tested in other populations (such as previously depressed outpatients) to assess the differential indication. A larger study will also allow to investigate the effects of the intervention, as well as moderators of the effect of mindfulness practice and text message feedback on clinical outcome.

Acknowledgments

This work is part of the “Competence Centre for the Prevention of Mental and Psychosomatic Disorders in Work and Educational Settings” (coordinators Prof Sabine Herpertz and Prof Wolfgang Herzog, Heidelberg University Hospital, Germany), funded by the Ministry of Science, Research and the Arts of the State of Baden-Württemberg, Germany. We are grateful to Lea Reichhardt and Miriam Birzle for supporting data collection and data entry, to Lutfi Arikan for technical setup and support of the automatic text message feedback system, and to the therapists of the study site, especially to Dr Karsten Tschauner, for supporting recruitment.

Conflicts of Interest

None declared.

References

1. Goyal M, Singh S, Sibinga EM, Gould NF, Rowland-Seymour A, Sharma R, et al. Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. JAMA Intern Med 2014 Mar;174(3):357-368 [FREE Full text] [doi: 10.1001/jamainternmed.2013.13018] [Medline: 24395196]
2. Strauss C, Cavanagh K, Oliver A, Pettman D. Mindfulness-based interventions for people diagnosed with a current episode of an anxiety or depressive disorder: a meta-analysis of randomised controlled trials. PLoS One 2014;9(4):e96110 [FREE Full text] [doi: 10.1371/journal.pone.0096110] [Medline: 24763812]
3. Chadwick P, Hughes S, Russell D, Russell I, Dagnan D. Mindfulness groups for distressing voices and paranoia: a replication and randomized feasibility trial. Behav Cogn Psychother 2009 Jul;37(4):403-412. [doi: 10.1017/S1352465809990166] [Medline: 19545481]
4. Didonna F. Mindfulness-based interventions in an inpatient setting. In: Didonna F, editor. Clinical handbook of mindfulness. New York: Springer; 2009:447-462.
5. Herdt J, Bührlen B, Bader K, Hänny C. Participation in an Adapted Version of MBCT in Psychiatric Care. Mindfulness 2012 Jun 27;3(3):218-226. [doi: 10.1007/s12671-012-0120-7]
6. Vettese LC, Toneatto T, Stea JN, Nguyen L, Wang JJ. Do mindfulness meditation participants do their homework? And does it make a difference? A review of the empirical evidence. J Cogn Psychother 2009 Aug 01;23(3):198-225. [doi: 10.1891/0889-8391.23.3.198]
7. Crane C, Crane RS, Eames C, Fennell MJ, Silverton S, Williams JM, et al. The effects of amount of home meditation practice in Mindfulness Based Cognitive Therapy on hazard of relapse to depression in the Staying Well after Depression Trial. Behav Res Ther 2014 Dec;63:17-24 [FREE Full text] [doi: 10.1016/j.brat.2014.08.015] [Medline: 25261599]
8. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. J Med Internet Res 2014;16(10):e222 [FREE Full text] [doi: 10.2196/jmir.3442] [Medline: 25326646]
9. Bauer S, Okon E, Meermann R, Kordy H. Technology-enhanced maintenance of treatment gains in eating disorders: efficacy of an intervention delivered via text messaging. J Consult Clin Psychol 2012 Aug;80(4):700-706. [doi: 10.1037/a0028030] [Medline: 22545736]
10. Shapiro J, Bauer S. Use of short message service (SMS)-based interventions to enhance low intensity CBT. In: Bennett-Levy J, Richards DA, Farrand P, Christensen H, Griffiths C, Kavanagh D, et al, editors. Oxford guide to low intensity CBT interventions. Oxford, England: Oxford University Press; 2010:281-286.
11. Aguilera A, Muñoz RF. Text messaging as an adjunct to CBT in low-income populations: a usability and feasibility pilot study. Prof Psychol Res Pr 2011 Dec 1;42(6):472-478 [FREE Full text] [doi: 10.1037/a0025499] [Medline: 25525292]
12. Kristjánsdóttir Ó, Fors EA, Eide E, Finset A, van Dulmen S, Wigers SH, et al. Written online situational feedback via mobile phone to support self-management of chronic widespread pain: a usability study of a Web-based intervention. BMC Musculoskelet Disord 2011;12:51 [FREE Full text] [doi: 10.1186/1471-2474-12-51] [Medline: 21352516]
13. 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]
14. Cavanagh K, Strauss C, Cicconi F, Griffiths N, Wyper A, Jones F. A randomised controlled trial of a brief online mindfulness-based intervention. Behav Res Ther 2013 Sep;51(9):573-578. [doi: 10.1016/j.brat.2013.06.003] [Medline: 23872699]
15. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the 'app' in happiness: a randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. J Happiness Stud 2014 Oct 29;1(17):163-185. [doi: 10.1007/s10902-014-9589-1]
16. Krusche A, Cyhlarova E, King S, Williams JM. Mindfulness online: a preliminary evaluation of the feasibility of a web-based mindfulness course and the impact on stress. BMJ Open 2012;2(3):e000803 [FREE Full text] [doi: 10.1136/bmjopen-2011-000803] [Medline: 22614170]
17. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. BMJ Open 2014;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]
18. Plaza I, Demarzo MM, Herrera-Mercadal P, García-Campayo J. Mindfulness-based mobile applications: literature review and analysis of current features. JMIR Mhealth Uhealth 2013;1(2):e24 [FREE Full text] [doi: 10.2196/mhealth.2733] [Medline: 25099314]
19. Segal Z, Williams J, Teasdale J. Mindfulness-based cognitive therapy for depression: a new approach to preventing relapse. New York: The Guilford Press; 2002.
20. 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] [Medline: 11556941]
21. Ehring T, Zetsche U, Weidacker K, Wahl K, Schönfeld S, Ehlers A. The Perseverative Thinking Questionnaire (PTQ): validation of a content-independent measure of repetitive negative thinking. J Behav Ther Exp Psychiatry 2011 Jun;42(2):225-232 [FREE Full text] [doi: 10.1016/j.jbtep.2010.12.003] [Medline: 21315886]
22. Walach H, Buchheld N, Buttenmüller V, Kleinknecht N, Schmidt S. Measuring mindfulness—the Freiburg Mindfulness Inventory (FMI). Pers Individ Dif 2006 Jun;40(8):1543-1555. [doi: 10.1016/j.paid.2005.11.025]

23. Hupfeld J, Ruffieux N. Validation of a German version of the Self-Compassion Scale (SCS-D) [Validierung einer deutschen Version der Self-Compassion Scale (SCS-D)]. Z Klin Psychol Psychother 2011 Apr;40(2):115-123. [doi: 10.1026/1616-3443/a000088]
24. Raes F, Pommier E, Neff KD, Van Gucht D. Construction and factorial validation of a short form of the Self-Compassion Scale. Clin Psychol Psychother 2011 Jun;18(3):250-255. [doi: 10.1002/cpp.702] [Medline: 21584907]
25. Kordy H, Backenstrass M, Hüsing J, Wolf M, Aulich K, Bürgy M, et al. Supportive monitoring and disease management through the internet: an internet-delivered intervention strategy for recurrent depression. Contemp Clin Trials 2013 Nov;36(2):327-337. [doi: 10.1016/j.cct.2013.08.005] [Medline: 23974036]
26. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j.2002.384.doc.x] [Medline: 15189396]
27. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011 May;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]
28. Wolf M, Kraft S, Tschauner K, Bauer S, Becker T, Puschner B. User activity in a mobile phone intervention to assist mindfulness exercises in people with depressive symptoms. Ment Health Prev 2016 Jun;4(2):57-62. [doi: 10.1016/j.mhp.2016.02.003]
29. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
30. van der Velden AM, Kuyken W, Wattar U, Crane C, Pallesen KJ, Dahlgaard J, et al. A systematic review of mechanisms of change in mindfulness-based cognitive therapy in the treatment of recurrent major depressive disorder. Clin Psychol Rev 2015 Apr;37:26-39 [FREE Full text] [doi: 10.1016/j.cpr.2015.02.001] [Medline: 25748559]

Abbreviations

FMI: Freiburg Mindfulness Inventory
MBCT: mindfulness-based cognitive therapy
mHealth: mobile health
PHQ-9: Patient Health Questionnaire-9
PTQ: Perseverative Thinking Questionnaire
SCS: Self-Compassion Scale
SMS: short message service

Edited by A Aguilera; submitted 14.12.16; peer-reviewed by A Birney, K Kannisto; comments to author 17.01.17; revised version received 15.02.17; accepted 27.02.17; published 02.05.17.

Please cite as:
Kraft S, Wolf M, Klein T, Becker T, Bauer S, Puschner B
Text Message Feedback to Support Mindfulness Practice in People With Depressive Symptoms: A Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2017;5(5):e59
URL: http://mhealth.jmir.org/2017/5/e59/
doi:10.2196/mhealth.7095
PMID:28465278

Original Paper

eTEST: Developing a Smart Home HIV Testing Kit that Enables Active, Real-Time Follow-Up and Referral After Testing

Tyler Wray[1], PhD; Philip A Chan[2], MD; Erik Simpanen[1], BS; Don Operario[1], PhD

[1]Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, United States

[2]Division of Infectious Diseases, Brown University Alpert Medical School, Providence, RI, United States

Corresponding Author:
Tyler Wray, PhD
Department of Behavioral and Social Sciences
Brown University School of Public Health
121 S Main Street
Box G-S121-5
Providence, RI, 02903
United States
Phone: 1 9134882714
Fax: 1 4018636697
Email: tyler_wray@brown.edu

Abstract

Background: Men who have sex with men (MSM) are the group at highest risk for contracting human immunodeficiency virus (HIV) in the United States, but many do not test as frequently as recommended. Home-based self-testing (HBST) for HIV holds promise for promoting regular testing among these individuals, but currently available HBSTs have limited follow-up options, providing only a 1-800 number that participants can call. Failure to actively conduct follow-up counseling and referrals after HBST use could result in delays in seeking confirmatory testing and care among users receiving reactive (preliminary positive) test results. HBST also fails to connect users who test negative with other prevention services that can reduce their future risk for HIV.

Objective: The aim of our study was to use qualitative research methods with high-risk MSM to inform development of a "smart" HBST kit. The kit utilizes existing Internet-of-Things (IoT) technologies to monitor HBST use in real-time and enable delivery of timely, active follow-up counseling and referrals over the phone.

Methods: In phase 1, individual interviews (n=10) explored how participants might use HBST and their views and preferences for conducting counseling and referral after HBST. Based on these perspectives, we developed a smartphone app (iOS, Android) that uses data from light sensors on Bluetooth low energy (BLE) beacons to monitor when HBST kits are opened, facilitating timely follow-up phone contact with users. In phase 2, a usability study conducted among high-risk MSM (n=10) examined the acceptability and feasibility of this system and provided user perspectives after using the system along with HBST.

Results: Phase 1 themes suggested that MSM preferred HBST, that most thought active follow-up after HBST would be valuable, and that doing so over the phone within 24 h after testing was preferable. Phase 2 results showed that the eTEST system successfully detected HBST use in nearly all cases. Participant perspectives also suggested that the timing, method (ie, phone call), and duration of follow-up were appropriate and helpful.

Conclusions: Using BLE beacons and a smartphone app to enable follow-up counseling and referral over the phone after HBST use is feasible and acceptable to high-risk MSM. Future research is needed to compare the effects of follow-up counseling on rates of repeat testing and receipt of referral services (eg, testing for sexually transmitted infections and initiation of preexposure prophylaxis) and to explore the acceptability of the eTEST system over longer periods of time.

(JMIR Mhealth Uhealth 2017;5(5):e62) doi:10.2196/mhealth.6491

KEYWORDS

HIV; smartphone; Internet; counseling; referral and consultation; qualitative research

Introduction

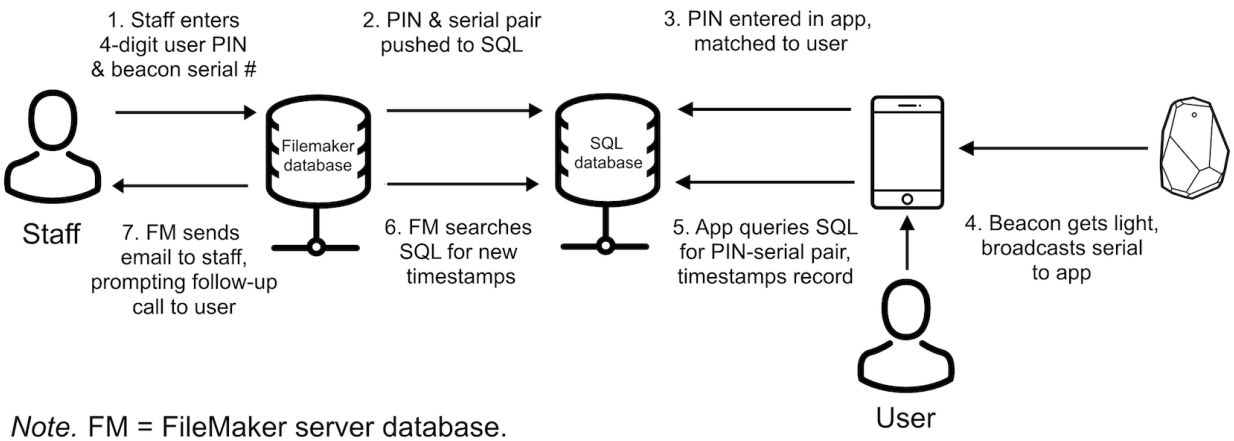
New human immunodeficiency virus (HIV) infections among men who have sex with men (MSM) in the United States continue to climb [1], with recent data showing that up to 1 in 6 MSM will be diagnosed with HIV in their lifetimes [2]. Modeling studies suggest that up to 50% of new HIV infections among MSM stem from the approximately 20% of those who are infected with HIV but are unaware of their status and thus are not virally suppressed [3-5]. These data highlight that an important step toward reducing HIV incidence in MSM involves increasing the accessibility and regularity of testing [6].

The first rapid antibody HIV test designed to be used and interpreted entirely by consumers at home was approved for over-the-counter sale by the US food and drug administration (FDA) in 2012 (OraSure Technologies, Bethlehem, PA; see Figure 1). The OraQuick home-based self-test (HBST) for HIV samples oral fluid and produces results in 20 min, and it has instructions that are simple and easy to understand. HBSTs could provide opportunities to expand and encourage regular HIV testing, especially among those who encounter barriers to testing at standard "brick-and-mortar" clinical sites. Research in developing countries using blood sample HBSTs has shown higher uptake of HBST compared with clinic-based testing [7-9]. Moreover, past studies in the United States show that many MSM, including those who have never tested, prefer to do so at home and believe they would test more often with HBST [10-15], since HBST addresses key barriers to clinic-based testing (eg, confidentiality and inconvenience [10,16]). Finally, past studies using popular gay-oriented social networking smartphone apps to connect high-risk MSM with free HBST showed that many users requested kits, and that they were useful for detecting new HIV infections among these users [17-20]. Moreover, 77% of these new cases were ultimately diagnosed at CD4 counts >350 cells/μL, suggesting that this approach may facilitate early diagnosis and treatment [20].

Despite their promise for overcoming barriers to testing, currently available HBSTs also have a number of important limitations. In some studies, the OraQuick test has shown lower sensitivity (95.9-99.6%) compared with other rapid antibody tests, especially when oral fluid is used [21,22]. The OraQuick also takes longer to detect antibodies after infection compared with fourth-generation and HIV RNA tests, meaning that some early or acute HIV infections could be missed [23]. However, some of these limitations may be unique to the OraQuick test; other rapid HIV tests with better sensitivity and the ability to detect more recent infections could be packaged for home use in the future, addressing some of these problems. Moreover, it may be best to use HBST as a compliment to these other, more precise tests (rather than a replacement), or as a method of engaging those at high-risk who would not otherwise test.

Another key limitation that is common to all existing HBSTs to date is the lack of posttesting follow-up and referrals. Many have argued that one of the most important benefits of clinic-based testing is that test counselors can personally link patients with reactive test results with confirmatory HIV testing and care, or refer those with negative results to other services that reduce HIV risk (eg, testing for other sexually transmitted infections [STIs], preexposure prophylaxis [PrEP], and risk reduction counseling) [24-26]. Modeling studies suggest that because of HBST's lack of follow-up and referral, widespread use of HBST may actually increase HIV incidence, since individuals who receive reactive results and are not linked with care may delay in seeking it, resulting in onward transmissions during that time [24,25]. Since the OraQuick test became commercially available, OraSure Technologies has maintained a 24-h, toll-free helpline that HBST users can call to receive instructions or guidance about how to conduct the test, posttest counseling, and referrals to HIV care [27]. However, this "passive" approach relies on HBST users to "reach out" for counseling, referrals, and linkage to care themselves, which may be insufficient for many at-risk MSM [10,24]. As of 2015, OraSure estimated that over 500,000 tests had been sold, but of 38,000 calls to their helpline, less than 5% of those were related to posttest counseling needs of HIV diagnosis or treatment [27]. Based on the expected rate of reactive results, these data lend support to concerns that many HBST users may not be connecting to follow-up and referral services after testing, including confirmatory testing or care, STI testing, and additional prevention resources.

Figure 1. Flow of eTEST system components.



Note. FM = FileMaker server database.

Conducting active follow-up with HBST users to provide these services after they test could overcome many of these limitations, while also helping high-risk MSM test more often. Similar to clinic-based testing, active follow-up after HBST could involve having trained paraprofessionals reach out to HBST users to link them with confirmatory testing and medical care, should their results be reactive. This would address one of the key priorities of posttest counseling. Active follow-up could also be beneficial for those with nonreactive results, since counselors could link users with other critical services for reducing their future HIV risk, such as STI testing, risk reduction counseling, safer sex supplies (eg, condoms, lube), PrEP, or postexposure prophylaxis (PEP). However, to be relevant to users, this follow-up contact must be timely, taking place soon after they use a HBST. While it would be difficult to conduct active follow-up with everyone who purchased a test commercially, making brief phone calls to particularly high-risk MSM who agree to receive test kits regularly in the mail might be more feasible. This could be an effective way of reaching a high-risk subset of MSM to encourage them to test more often, while also providing them with the essential posttest services they need (eg, risk reduction counseling, referrals for STI-testing or safe sex supplies, linking those with reactive results with care as soon as possible).

Given this need, we used qualitative research methods to iteratively guide the development and initial evaluation of a "smart" home-based HIV self-testing kit (ie, eTEST) that monitors when kits have been opened in real-time. This system allowed us to actively reach out to users after they use a HBST to provide follow-up counseling and referral over the phone. To inform the development of the eTEST system, phase 1 of this study explored how high-risk MSM would use HBST and their perspectives about receiving counseling and referral after HBST. In phase 2, we then examined the feasibility and acceptability of this system with another small sample of high-risk MSM after using "smart" HBST kits at home.

Methods

Participants

A total of 20 participants (n=10 for phase 1, n=10 for phase 2) were recruited via websites and smartphone apps that are commonly used by MSM to meet partners (eg, Grindr, Scruff, and so on) [28]. Eligible participants were (1) 18+ years old, (2) assigned male sex at birth, and (3) fluent in English. They also reported (4) not having been tested for HIV in the last year, (5) having had anal sex without a condom or without having taken PrEP with a casual male partner at least once in the preceding 6 months, and (6) had sex with a casual male partner met on the Web in the past year. Participants were also required to have (7) a stable residence and (8) an iOS or Android (version 4.3 or higher) smartphone with a data plan. The same eligibility criteria applied to participants enrolling in either phases 1 or 2, and participants were required to meet all criteria. All participants reported being male gender, and no participants reported having taken an HBST, PrEP, or PEP in the past.

Phase 1: Preliminary Interviews

Phase 1 interviews were conducted individually at our offices. The main focus of this phase was to understand HIV testing and how these participants might use HBST, so as to inform our approach to developing the eTEST program, as well as the contexts in which HBST might be used. Key questions posed in this phase included how frequently participants do or think they should test for HIV, their personal barriers to testing, preferences for HBST versus clinic-based testing, views about offering active follow-up after HBST, and preferences about how follow-up might be provided. Whereas the majority of participants in this phase (and phase 2) reported never having tested for HIV before, all had heard about the process and commented on their perceptions based on what they knew about HIV testing. These interviews lasted an hour, were tape recorded (for later transcription and analysis), and participants were paid US $50.

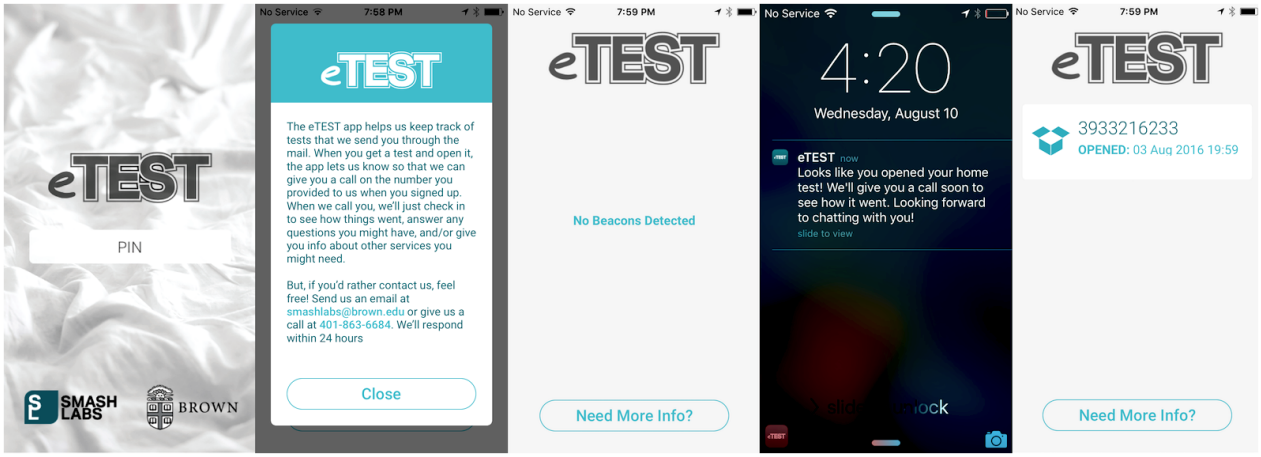
The eTEST System: Monitoring Approach

After analyzing and interpreting data from phase 1, we used these data to guide the development of the eTEST system. The system used a smartphone app (both iOS and Android versions), together with newly available technologies often used for "Internet-of-Things" (IoT) applications called Bluetooth low energy (BLE) beacons, to monitor when specific users opened a HBST kit that had been sent to them in the mail. BLE beacons are small electronic devices that broadcast a radio signal that can be received by any device equipped with Bluetooth. Whereas BLE beacons are most commonly used to improve smartphone location while indoors (see kontakt.io.Inc., for more information [29]), beacons can also be fit with a variety of sensors (temperature, light, or motion), so that these data can be broadcast to Bluetooth-enabled devices as well.

To set up users in the system, staff entered the user's information into a secure, staff-side database (Filemaker server). This assigned each user a unique 4-digit personal identification number (PIN), and "pushed" these data to a structured query language (SQL) database that interacted directly with the eTEST mobile app (Figure 1).

Staff then guided users through downloading the eTEST app onto their smartphones and entered their assigned 4-digit PIN number in the app, which matched it with the PIN entered through the staff-side database (unmatched PINs produced an error). This way, the app did not collect or store any personal or identifying data, only a participant's unique study ID (PIN). After successful initialization, the app displayed basic information to the user about the app's purpose, as well as contact information for study staff (Figure 2).

Figure 2. eTEST smartphone app screens.



Users could access this information at any time by pressing a button from the app's home screen. When preparing an HBST test kit to be mailed, staff fit a BLE beacon to the lid of the plastic test enclosure and logged the serial number of the beacon into a specific user's record in the staff-side database, which pushed this PIN-serial combination to the SQL database (see Multimedia Appendix 1). To monitor whether a test kit has been opened, the eTEST app used data from the BLE beacon's ambient light sensor and registered the kit as "open" when the level of ambient light reaching the beacon was >3 lux. As such, when beacons came within range of a user's smartphone (≈50m) and received sufficient light (> 3 lux), the app pushed a notification to the user and relayed these data to the SQL database, entering a timestamp for the specific PIN-serial combination detected. This identified the user and test kit that was opened in the staff-side database and prompted an automated email to study counselors that notified them of the need to place a call to a user within 24 h of receipt. These notifications were sent only once, when the user first opened the test kit.

We selected Estimote Location beacons for this project, which were released in March of 2016. These beacons required no special programming and only minimal set up to use with the eTEST app. The set up was easily accomplished by staff members and typically took less than a minute (see Multimedia Appendix 2). In addition to the integrated ambient light sensor, these beacons also had a "dark to sleep" feature, that tells the beacon to stop broadcasting data until it receives a level of light that is above a customizable threshold. Using this feature, beacons can be effectively "muted" until the beacon receives sufficient light (>3 lux). This capability was important for our use case, since our goal in designing the eTEST system was to ensure that it could detect HBST use over long periods of time without requiring any intervention from users (other than initially downloading the app). That is, we wanted to design the eTEST system to successfully capture HBST kit openings, even when users had the app running in the background, when they had "killed" the app, or when the phone was in sleep mode. Accomplishing this proved to be a somewhat difficult technological hurdle in iOS. With a typical BLE beacon that continuously broadcasts data, apps that have been "killed" can be "woken up" when a beacon comes within range (typically ≈50m) of the phone. However, iOS kills these processes again after a few minutes. This posed a problem for our use case, since the beacon (enclosed in a test kit) would often be expected to come within range of a user's phone long before they actually open the test kit. That is, the time between users receiving a test kit and when they actually open it will often be significant. So, running the app's processes briefly only after users receive the test posed a high risk for missing opening events, if iPhone users had killed the eTEST app or if the phone was in sleep mode. Using the "dark to sleep" feature of the Estimote Location beacons, we directed the beacons to begin broadcasting only after it received a sufficient amount of light, prompting the app to briefly "wake up" to sync with the server. Preliminary testing in the lab showed that, using this feature, test kit openings were indeed successfully detected in virtually all iPhone states (eg, app "killed," phone in sleep mode), so long as the app had been downloaded and had been successfully paired with a 4-digit PIN and beacon serial number. Of note, this limitation is not relevant to Android users, since a process can be run in the background at any time.

Phase 2: Usability Testing and Follow-Up Interviews

After initially developing the eTEST system, we then iteratively built several app releases and tested them with various devices and states, both in the lab and during a number of "hallway tests" [30]. Once a working version was constructed, we moved on to phase 2, which involved conducting a usability test with target users [31]. In this phase, our goal was to preliminarily test how well the eTEST monitoring and follow-up system worked in a real-world scenario, where participants were using their phones, the app, and the HBST kits under typical circumstances. We explored several questions that could affect how well the system works, including whether users opened the "smart" HBST kit within range of their smartphones, whether they actually used the test soon after initially opening it, whether the light threshold (>3 lux) was sufficient to detect openings in dimly-lit environments, and whether detection was successful across operating systems (ie, iOS and Android) and phone or app states (eg, app "killed" or phone in sleep mode). We also examined ease of use, perceived utility of (and comfort with) receiving follow-up calls, the length and timing of calls, and concerns about confidentiality or privacy.

To address these questions, participants in phase 2 first met with research staff in-person to learn about the study, downloaded the eTEST app onto their smartphones, and were provided with a BLE-enabled OraQuick home HIV test. They were asked to use the test at some point within 7 days, but were provided with no other instructions about how to use the app, other than to keep it downloaded onto their device during that time. This allowed us to explore how well the eTEST system worked in a variety of phone states. After they opened the test, a qualified HIV test counselor (QHTC) followed up with participants over the phone within 24 h. During these calls, QHTCs provided a "typical" sequence of posttest counseling, including: (1) discussion of the test, its use, and results; (2) offering counseling to reduce HIV risk behaviors; and (3) offering referrals for STI testing, safe sex information and supplies, PrEP consultation, and other medical care or counseling (mental health, alcohol or drug). Afterward, counselors conducted a brief follow-up qualitative interview about their experience and opinions of eTEST. These interviews lasted about 30 min, were recorded, and participants were compensated $50. All procedures were approved by the Brown University Institutional Review Board.

Analysis

All interviews and counseling calls were transcribed, and transcripts were reviewed and coded manually by study staff. Content and codes were then analyzed thematically. Regular discussions helped to arrive at themes that emerged across participants, and specific quotes were excerpted from transcripts to illustrate these themes.

Results

Phase 1: Preliminary Interviews

Ten participants completed the phase 1 interview. Only 3 of these participants (30%, 3/10) reported having tested for HIV in their lifetime, and of those, all reported having tested only at a clinic or outreach event (see Table 1).

However, all participants thought that MSM in general should be tested more often than they tested themselves: whereas no interviewees reported having been tested in the past year (since this was a criteria for eligibility), most suggested that MSM should be tested at least every six months, and with more frequent testing for those who engaged in riskier sexual behavior. Those that had tested before noted that they appreciated testing in person with a paraprofessional because the counselors were often understanding, skilled, and could make testers feel comfortable. They also noted that in-person HIV tests were also relatively quick and easy to do. However, most participants noted drawbacks of in-person HIV testing, with most identifying nervousness as a key downside:

> *Basically throughout the entire duration of the tests or going into the clinic up to getting your results, I felt nervous about what the results might be, even though I felt for my case that it was extremely unlikely that something would come back positive.* [Participant 1, 30-year-old white male, last tested 3.2 years ago]

Despite the professionalism of clinic staff, there were also concerns about the confidentiality of clinic-based testing:

> *The confidentiality. That's a big thing (...) because I know people who work in there—I was afraid of like the rumor getting around.* [Participant 2, 21-year-old Hispanic male, last tested 1.1 years ago]

Barriers to More Frequent Testing

Most participants noted that the most important barrier was that they perceived themselves to be at low risk, despite meeting the study inclusion criteria. However, another key theme mentioned by most participants was the inconvenience involved:

> *It's something that I actually have to go out and do, and it just kind of slips my mind a lot, more than anything. And then it's like I know I should, and then I don't know, I'm really good at procrastinating.* [Participant 3, 36-year-old white male, last tested 2.1 years ago]

This concept illustrates the importance of making testing as quick and easy to do as possible. By delivering free HBST kits directly to a subset of high-risk individuals at recommended intervals (eg, at least once every six months), the convenience of testing at home may help overcome other important barriers like low risk perceptions.

Pros and Cons of Home-Based Self-Testing (HBST) and Preferences for HBST Versus Clinic-Based Testing

Most phase 1 participants noted the convenience of HBST as its biggest strength:

> *A doctor is not always going to be in the office when you call. Sometimes it's hard to reach a doctor. It could take weeks. But somebody (who) buys this doesn't have to wait that long at all. The same day they buy it, they can get results the same day.* [Participant 4, 42-year-old Hispanic male, never tested]

Most participants also noted that HBST could be more confidential and private than clinic-based testing:

> *It's extremely private...nobody will know the results but you. It's a good way to be more private or at ease for certain people.* [Participant 1, 30-year-old white male, last tested 3.2 years ago]

Finally, several participants also mentioned that doing the HIV test themselves might make them feel more empowered and proud of having taken responsibility for their own health:

> *I think it gives the individual more power and control over the situation. In the past, you had to go to a doctor in order to get tested. But now, the individual has the power to access that knowledge.* [Participant 3, 36-year-old white male, last tested 2.1 years ago]

Table 1. Participant demographic characteristics.

Characteristics	Median (IQR[a]) or n (%)	
	Phase 1 (n=10)	Phase 2 (n=10)
Age (years), range: 21-67, mean (SD)	29 (18)	30 (31)
Race		
White	8 (80)	9 (90)
Black or African American	2 (20)	1 (10)
Ethnicity (Hispanic or Latino)	2 (20)	0 (0)
Relationship status		
Single or never married	8 (80)	7 (70)
In a committed relationship	0 (0)	0 (0)
In a domestic partnership	0 (0)	0 (0)
Married	0 (0)	1 (10)
Separated	2 (20)	0 (0)
Divorced	0 (0)	1 (10)
Widowed	0 (0)	1 (10)
Education		
High school diploma or general educational development	1 (10)	1 (10)
Some college education	6 (60)	4 (40)
College graduate	2 (20)	3 (30)
Graduate or professional degree	1 (10)	2 (20)
Income		
$0-$29,999	5 (50)	3 (30)
$30,000-$99,999	5 (50)	6 (60)
$100,000 or more	0 (0)	1 (10)
Sexual identity		
Gay	9 (90)	8 (80)
Bisexual	0 (0)	1 (10)
Other	1 (10)	1 (10)
Never tested for HIV	7 (70)	9 (90)
Average years since last HIV test[b], mean (SD)	2 (2.1)	16 (--)

[a]IQR: interquartile range

[b]Among participants reporting having tested for HIV in their lifetimes.

In addition to these strengths, however, interviewees also noted a number of important drawbacks of HBST compared with more traditional, clinic-based testing. Many target users noted a fear of conducting the test incorrectly when doing it on their own as a key reservation:

> *I would be super terrified that I would be messing something up, and I wouldn't actually know how to use it, so I don't know if there's a way to somehow make a YouTube video or something or have (...) like a Q-R code that went to a video that showed how it's done. As a millennial, I don't like to read. Things need to be like quick and easy.* [Participant 5, 21-year-old white male, never tested]

Participants also noted concerns about the potential consequences of getting results from an HIV test while alone as another drawback of HBST:

> *What happens when the result comes back positive? With a clinic or some other kind of professional present, there's someone who can handle the situation in some ways if I'm not able to myself. And here, someone who is not able to handle the situation, for example by calling a professional or going to the doctor afterwards, would be left alone.* [Participant 3, 36-year-old white male, last tested 2.1 years ago]

Finally, when asked whether they would prefer using a HBST kit versus testing at a clinic, participants overwhelmingly said they would prefer to use HBST:

> *You can't say 'Oh, I don't want to go make that doctor's appointment when you've got this sitting in your cabinet and you can use it. (...) I can't think of too many cases where you wouldn't just want to do (HBST) to save yourself the time, hassle, and stigma (...) of going to the clinic.* [Participant 8, 48-year-old African American male, never tested]

Perceptions of Offering Follow-Up after HBST

When we asked these participants about their views of offering more active follow-up counseling and referrals after HBST by reaching out to users (as opposed to providing users with the phone number of a 24-h helpline), most thought this would be helpful. The advantage most commonly identified by participants was that actively following up with HBST users would show concern and provide support:

> *There's something always great about talking to another human being. (...) You're just left kind of very vulnerable out there, and you don't have the structure around you that you would if you went to a medical clinic, and that sense of direction and extra support from people to come in and help you would be nice. It's bringing in that support structure that's already in the medical clinic into the home.* [Participant 2, 21-year-old Hispanic male, last tested 1.1 years ago]

Participants also suggested that this active follow-up could be especially important for users who receive reactive results through a HBST kit, specifically because it could help link them with confirmatory testing and follow-up care:

> *I suppose like if you did find your result to be positive that could be kind of shocking and that you might really not be sure what to do. It would be an advantage to have someone there with you to talk about what to do.* [Participant 3, 36-year-old white male, last tested 2.1 years ago]

Another commonly identified advantage was that having a counselor reach out could encourage users to follow up with referrals and test again in the future:

> *I think the reaching out, it shows concern. I think it would spark enthusiasm to get serious about (getting tested).* [Participant 7, 42-year-old African American male, never tested]

A few participants thought offering active follow-up was either unnecessary because users could find information by themselves, or because it could deter some users from testing who would prefer not to speak with someone afterward:

> *I might want additional information or referrals, and assuming the test doesn't offer it, then I would be able to get that. But one of the biggest advantages of this test that I see is what people who are afraid of talking to a doctor or a person about HIV could use it. And I would rather have those people be able to use the test than not take it at all.* [Participant 9, 24-year-old white male, never tested]

Finally, participants overwhelmingly believed that receiving a phone call would be the best way for counselors to follow up after using an HBST kit. However, others suggested they would be equally comfortable with text-based forms of communication, like chat, text, or even email:

> *I think (calling is) very old, outdated—people have changed, technology has changed. For me, I guess I respond best by email or texting. Maybe if there was an app, that would be super cool.* [Participant 5, 21-year-old white male, never tested]

Phase 2: Usability Testing

Ten participants completed phase 2, and of these, only 1 (10%, 1/10) reported having ever tested for HIV. No participants reported having taken or heard of HBST prior to participating. Seven participants used iOS smartphones and 3 used Android smartphones. All users in phase 2 reported having tested negative.

Nine of ten opening events were successfully captured, and by participant report, notifications typically registered within a minute after having opened their HBST kits. Follow-up calls were successfully placed within 24 h to each of the 9 users for whom the system worked as expected. Follow-up interview data with the remaining user suggested that the system failed because his smartphone was being repaired when he opened the test. Together, the results of our lab testing and this usability test show that the eTEST system was able to successfully detect when "smart" HBST kits were opened by users, except in less common circumstances in which participants' phones were not functional.

Phase 2 qualitative interviews also suggested that our strategy of monitoring when users opened their tests using a BLE beacon and smartphone app was a good fit with how participants used their tests at home. For example, one key question about this approach was whether users actually took their home-based tests soon after initially opening the test enclosure, or whether there would be a delay, making efforts to follow up after initially opening the test less relevant. However, results suggest that all users took their tests within a few minutes after opening them. Since the smartphone can only detect data from beacons within a certain proximity (≈50 m), another question was about whether users would have their phones nearby them when they took the test. All participants reported having their phones within a few feet of the test when opening it, with some adding that they used other features of their phones to help them with the test (eg, the "timer" app). Finally, users also noted that the eTEST app did not appear to drain their phone battery noticeably, and that it used very minimal data.

User Perspectives

Feedback from these participants about home-based testing and the eTEST system was very positive. Participants emphasized the simplicity of the home-based test itself, with several also noting that providing the test kits in the mail at regular intervals may encourage them to test more regularly:

> *It was just so easy to follow the directions (of the HBST). They seemed to think of everything. I kind of enjoyed doing it at home, it seemed very straight-forward. If that thing came in the mail on a schedule, I think somebody would be much more likely to do it, as opposed to like 'Oh, I've got to go down there) to get tested).* [Participant 17, 22-year-old white male, never tested]

Many participants also noted that having QHTCs follow up with them after they tested over the phone could provide important support, and may help many overcome obstacles they have in seeking help afterward:

> *I find it unnerving in these kinds of circumstances to take the first step of getting that information, so to have someone else there saying 'Oh, I have this information for you right now,' I think is very helpful, because sometimes (finding) those things on your own can be a bit stressful.* [Participant 11, 22-year-old white male, never tested]

Several also noted that following up with users after they test at home could be especially important for those who have reactive results (hypothetically), and could play a role in motivating these users to seek timely confirmatory testing and follow-up care:

> *I think it's good to have a follow-up, in case there was a situation where it was positive, I would know that someone would be contacting me to give me clinical information and places I could call to set up an appointment or get in contact to continue the process of getting a diagnosis. It definitely would force you to continue with a follow-up.* [Participant 13, 27-year-old white male, never tested]

Finally, some users said that, even for those who test negative, follow-up calls could be helpful for connecting users with referrals for other services (STI testing, in particular):

> *Most of the advertising you see is for HIV (testing), so I'd never heard of a place or a clinic you could go to have (STI testing) done until now.* [Participant 11, 22-year-old white male, never tested]

Participants also provided feedback about the method of offering counseling and follow-up after testing (ie, phone calls), as well as the timing and duration of the calls (which lasted on average 10 min; range 5-19 min). Most participants preferred to receive counseling and follow-up over the phone, but a few noted having other options to seek information and ask questions might be a better fit for some users:

> *The phone, I imagine, would be the simplest for a lot of follow-up conversations. The questions back and forth are easier to answer than maybe like an email would be. With the way younger people are used to technology these days, though, it's more about text messaging and something that's not as personal as actually seeing somebody.* [Participant 14, 62-year-old white male, never tested]

All participants agreed that the duration of the calls was appropriate, and that the topics addressed during the follow-up were relevant and useful. However, they disagreed about the timing of calls. Nearly all suggested that placing these calls within 24 h would ensure that the conversation would be most relevant; however, some participants said they would have preferred getting the call even sooner. Several noted that the best time to call might depend on the test results, and that, if they had received reactive results, they may want to be contacted sooner:

> *(If the results were reactive,) my gut says within an hour, I would want the phone to ring. If it was too long, who knows what I might do? It's so treatable that I highly doubt (anything would happen), but I think I would still want the call.* [Participant 20, 45-year-old white male, never tested]

Discussion

Principal Findings

Our goal in developing the eTEST system was to make testing for HIV more accessible, especially for high-risk MSM who encounter barriers to clinic-based testing services. HBSTs for HIV could be a key opportunity to help achieve this, since they eliminate some of the most common barriers that prevent high-risk individuals from testing more regularly in clinics [10,16,32]. One way to encourage more regular testing might be to deliver HBST to those at highest risk at specific intervals (eg, 3 months or 6 months), so that new infections can be detected earlier [24,25]. However, to ensure that those who use HBST can be efficiently connected with vital posttesting resources (eg, confirmatory testing and care), a tool that can monitor HBST use in real time and facilitate post-test follow-up and referral is needed. Results from this study provide strong support for the feasibility and acceptability of a tool we developed to accomplish this.

Consistent with past research [11,12,32], our preliminary interviews with high-risk MSM who met partners on the Web and who had not tested for HIV in the last year suggest that access to HBST may indeed encourage these individuals to test more regularly, and that most viewed HBST as more convenient, confidential, private, and empowering. Nearly all participants in both phases also reported that they believed they would be more likely to test using HBST than getting tested at a clinic. However, another important goal of developing the eTEST system was to address one of the key limitations of HBST: The lack of follow-up and linkage to care after testing at home. The results of our initial interviews suggested that, whereas some high-risk MSM may not prefer it, most saw advantages of offering counseling and referral over the phone after using a HBST. Specifically, participants noted that phone calls from an HIV test counselor would show concern, provide reassurance, and could connect them with professional help. Participants who actually experienced this sequence of follow-up shared these perspectives. Many also discussed ways to reduce their risk with counselors during the calls and received referrals for STI testing and PrEP.

Usability study findings provide key preliminary support for the feasibility and acceptability of our approach to monitoring participants' use of HBST kits (thereby enabling timely, active

follow-up phone calls). In developing the system, we hoped to ensure that it would work without requiring users to interact with their phones, since we expected that most would-be users of this system will not keep the app running or have their phones actively awake or in-use at the time they test. This is particularly important, given that we designed this system to be used over long periods of time, since even high-risk MSM would be encouraged to test once every 3 months, at most. With our approach, we were able to successfully detect 90% (n=9) of test kit opening events across both Android and iOS smartphones, with the one "missed" event occurring due to the user having his phone repaired. Also key to long-term acceptability was ensuring that the eTEST app was not overly burdensome and did not interfere with users' day-to-day use of their phones. All usability study participants kept the eTEST app on their phones for the week required and reported that the app did not noticeably affect their phone's battery, data usage, or memory (iOS: 33.4 MB, Android: 20.0 MB). Similarly, usability testing also confirmed that our strategy for monitoring HBST kits addressed many of the parameters needed for the system to work successfully. Specifically, we found that users took their tests soon after opening the test's plastic enclosure, suggesting that monitoring when the test kit is opened serves as an appropriate marker of when they actually took the test. Results also confirmed that participants often had their phones very close to them when they completed the test, with some using other features of their phones to help them with the test (eg, to time the required 20-min wait before reading results). This was important to confirm, since users' phones must be within approximately 50 m of the beacons in the test kits to capture data being transmitted from them.

Perspectives from usability study participants also provided support for the utility, duration, and timing of follow-up counseling and referral phone calls. Echoing phase 1 results, several usability study participants reported that receiving the follow-up calls provided them with support and reassurance, even though all test results were negative. Many also noted that the counselor's ability to provide them with further information about HIV, the test, and other prevention services was a strength. Finally, usability participants also noted that follow-up calls were timely and brief. Some reported that it might be ideal to connect with users who receive reactive results sooner than 24 h later; however, most said that the timing seemed appropriate and that the duration was brief enough to encourage them to take the call. Finally, participants also noted that they had no concerns about the privacy of the app or its features (eg, push notifications), since all register within a few minutes of opening their tests, ensuring that they were already in a private location when it displayed. Additionally, most had already made use of their devices' lock screen feature. Overall, these perspectives suggest that the HBST kits were easy to use and that the follow-up phone calls could be an important way to offer support to HBST users in general. They also suggest that reaching out to users afterward could be an important way to link them with other important resources like STI testing and PrEP, and with confirmatory testing and care (for those who receive reactive results).

Limitations

Several important limitations should be noted. First, as this system relies on a smartphone app that uses data, it can only be used by those with iOS or Android (version 4.3 or higher) smartphones that have a data plan. Market research suggests that up to 68% of the adult population in the United States currently own smartphones [31], and this figure is likely to grow exponentially in the future. iOS and Android users comprise 95.3% of these smartphones [32]. As demonstrated in our usability study, the eTEST system could be used to target MSM who use smartphone apps to meet partners on the Web, and therefore already use smartphones. However, some of those at highest risk for HIV may not own or use smartphones [33,34]. Similarly, the eTEST system was designed with the future goal of facilitating a free HBST program that involves regularly sending kits to participants in the mail. As such, its utility among particularly vulnerable and at-risk individuals, such as the homeless, may be limited. However, future research could explore alternative methods of test delivery and detection that are more fitting for individuals with unstable housing. Second, several characteristics of the technologies used may lead to "missed" test opens. Opening events are successfully captured even when the app is not running (even in the background) and the phone is in sleep mode; however, the phone must be turned on and the app must be on the phone and registered (ie, successfully paired with a 4-digit PIN number) at the time the test is opened. In addition, the phone must be within about 50 meters of the beacon. Finally, it should be noted that, while we are excited about the implications this work may have for improving the delivery of HBST, we do not believe HBST should replace testing at a clinic. More research is needed, but we believe the most fitting approach might involve encouraging HBST use in between visits to a clinic to get tested.

Future Research

Whereas these data show that the system operates as intended and that the perspectives and preferences of a small pool of intended users (ie, high-risk MSM) are supportive, additional research is needed to address a number of key questions about the utility of offering HBST programs to high-risk groups. Specifically, a key next step should involve testing whether HBST-based strategies improve rates and frequencies of HIV testing among high-risk populations that typically test infrequently, and in particular, whether these strategies are capable of reducing incidence among them as a result. To address this, future research should explore whether HBST approaches detect more new infections and facilitate earlier diagnoses compared with relying on clinic-based testing alone, as well as whether the eTEST approach is successful in linking users with key additional services that can prevent or reduce the risk of future infections (eg, confirmatory testing and care, STI testing, PrEP). With these data, researchers can model the potential population-level effects of incorporating HBST-based strategies into community testing programs on HIV incidence. Arguably the most important benefit of conducting posttest follow-up after HBST is the potential it has for linking those who receive reactive results with confirmatory testing and care as soon as possible. This could suggest that it is best to focus follow-up efforts toward these users in some way; for example,

by using the eTEST system to generate an automated SMS text message asking participants to report the results of their test, and then placing follow-up calls to only those who report receiving a reactive result. However, the lack of follow-up and referral to other prevention services (eg, STI testing, PrEP) for those who test negative has also been cited as another key limitation of HBST [24,25]. Linking users with these services may also provide valuable benefits by reducing future risk. As such, we believe that the question of "how" to offer follow-up and "with which users" is a question best left to future studies. For this reason, this manuscript details the rationale and approach for offering follow-up that could be useful for both types of users.

In summary, this study illustrates our use of emerging IoT technologies to enable more robust delivery of counseling and referral services alongside HBST for HIV. Iterative, qualitative research provided preliminary support for the feasibility and acceptability of using this approach to monitor when HBST users take their tests in order to facilitate timely follow-up phone calls from counselors. We are excited about the potential these technologies have for encouraging high-risk individuals to test more frequently; however, further work is needed to explore acceptability among target users over broader time periods (eg, months and years) as well as the effects that frequent HBST and follow-up phone contact have on rates of HIV testing and linkage to other prevention resources.

Acknowledgments

This manuscript was supported by R21MH109374 from the National Institute on Mental Health and L30AA023336 from the National Institute on Alcohol Abuse and Alcoholism. PAC is supported by the National Institute of Allergy and Infectious Diseases (K23AI096923).

Conflicts of Interest

None declared.

Multimedia Appendix 1

The OraQuick home-based self-test for HIV, fit with Estimote Location beacon.

[JPG File, 1MB - mhealth_v5i5e62_app1.jpg]

Multimedia Appendix 2

The Estimote Location Bluetooth low energy (BLE) beacon.

[JPG File, 518KB - mhealth_v5i5e62_app2.jpg]

References

1. Centers for Disease Control and Prevention. Atlanta, GA: U.S. Department of Health and Human Services; 2016. Trends in HIV diagnoses, 2005-2014 URL: https://www.cdc.gov/nchhstp/newsroom/docs/factsheets/hiv-data-trends-fact-sheet-508.pdf [accessed 2017-04-21] [WebCite Cache ID 6ptM40TFS]
2. Centers for Disease Control and Prevention. Redproject. Atlanta, GA: National Center for HIV/AIDS, Viral Hepatitis, STD and TB Prevention; 2016. Lifetime risk of HIV diagnosis URL: http://redproject.org/wp-content/uploads/2015/04/Lifetime-Risk-of-HIV-Diagnosis-in-the-United-States_2016.pdf [accessed 2017-04-21] [WebCite Cache ID 6ptMGosrp]
3. Centers for Disease Control and Prevention. Atlanta, GA: U.S. Department of Health and Human Services; 2015. National HIV Prevention Progress Report, 2015 URL: https://www.cdc.gov/hiv/pdf/policies/progressreports/cdc-hiv-nationalprogressreport.pdf [accessed 2017-04-21] [WebCite Cache ID 6ptMOn4yH]
4. Hall HI, Holtgrave DR, Maulsby C. HIV transmission rates from persons living with HIV who are aware and unaware of their infection. AIDS 2012 Apr 24;26(7):893-896. [doi: 10.1097/QAD.0b013e328351f73f] [Medline: 22313960]
5. Wilson DP, Hoare A, Regan DG, Law MG. Importance of promoting HIV testing for preventing secondary transmissions: modelling the Australian HIV epidemic among men who have sex with men. Sex Health 2009 Mar;6(1):19-33. [Medline: 19254488]
6. Granich R, Gilks C, Dye C, De Cock KM, Williams B. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 2009 Jan 03;373(9657):48-57. [doi: 10.1016/S0140-6736(08)61697-9] [Medline: 19038438]
7. Fylkesnes K, Siziya S. A randomized trial on acceptability of voluntary HIV counselling and testing. Trop Med Int Health 2004 May;9(5):566-572 [FREE Full text] [doi: 10.1111/j.1365-3156.2004.01231.x] [Medline: 15117300]
8. Matovu JK, Kigozi G, Nalugoda F, Wabwire-Mangen F, Gray RH. The Rakai Project counselling programme experience. Trop Med Int Health 2002 Dec;7(12):1064-1067 [FREE Full text] [Medline: 12460398]

9. Wolff B, Nyanzi B, Katongole G, Ssesanga D, Ruberantwari A, Whitworth J. Evaluation of a home-based voluntary counselling and testing intervention in rural Uganda. Health Policy Plan 2005 Mar;20(2):109-116. [doi: 10.1093/heapol/czi013] [Medline: 15746219]
10. Bilardi JE, Walker S, Read T, Prestage G, Chen MY, Guy R, et al. Gay and bisexual men's views on rapid self-testing for HIV. AIDS Behav 2013 Jul;17(6):2093-2099. [doi: 10.1007/s10461-012-0395-7] [Medline: 23297083]
11. Carballo-Diéguez A, Frasca T, Dolezal C, Balan I. Will gay and bisexually active men at high risk of infection use over-the-counter rapid HIV tests to screen sexual partners? J Sex Res 2012;49(4):379-387 [FREE Full text] [doi: 10.1080/00224499.2011.647117] [Medline: 22293029]
12. Mutale W, Michelo C, Jürgensen M, Fylkesnes K. Home-based voluntary HIV counselling and testing found highly acceptable and to reduce inequalities. BMC Public Health 2010 Jun 17;10:347 [FREE Full text] [doi: 10.1186/1471-2458-10-347] [Medline: 20553631]
13. Myers JE, El-Sadr WM, Zerbe A, Branson BM. Rapid HIV self-testing: long in coming but opportunities beckon. AIDS 2013 Jul 17;27(11):1687-1695. [doi: 10.1097/QAD.0b013e32835fd7a0] [Medline: 23807269]
14. Spielberg F, Branson BM, Goldbaum GM, Lockhart D, Kurth A, Celum CL, et al. Overcoming barriers to HIV testing: preferences for new strategies among clients of a needle exchange, a sexually transmitted disease clinic, and sex venues for men who have sex with men. J Acquir Immune Defic Syndr 2003 Mar 01;32(3):318-327. [Medline: 12626893]
15. Myers J. HIV self-testing in New York City. 2015 Feb 23 Presented at: Conference on Retroviruses and Opportunistic Infections; February 23, 2015; Seattle, WA.
16. Mackellar DA, Hou S, Whalen CC, Samuelsen K, Sanchez T, Smith A, et al. Reasons for not HIV testing, testing intentions, and potential use of an over-the-counter rapid HIV test in an internet sample of men who have sex with men who have never tested for HIV. Sex Transm Dis 2011 May;38(5):419-428. [doi: 10.1097/OLQ.0b013e31820369dd] [Medline: 21183863]
17. Rosengren A, Huang E, Daniels J, Young S, Marlin R, Klausner J. Feasibility of using Grindr(TM) to distribute HIV self-test kits to men who have sex with men in Los Angeles, California. Sex Health 2016 May 23 Epub ahead of print. [doi: 10.1071/SH15236] [Medline: 27209076]
18. Klausner J. Innovations in HIV self-testing delivery: Vending machines, vouchers, and US mail. 2015 Feb 23 Presented at: Conference on Retroviruses and Opportunistic Infections; February 23, 2015; Seattle, WA.
19. Huang E, Marlin RW, Young SD, Medline A, Klausner JD. Using Grindr, a smartphone social-networking application, to increase HIV self-testing among Black and Latino men who have sex with men in Los Angeles, 2014. AIDS Educ Prev 2016 Aug;28(4):341-350. [doi: 10.1521/aeap.2016.28.4.341] [Medline: 27427928]
20. Elliot E, Rossi M, McCormack S, McOwan A. Identifying undiagnosed HIV in men who have sex with men (MSM) by offering HIV home sampling via online gay social media: a service evaluation. Sex Transm Infect 2016 Sep;92(6):470-473. [doi: 10.1136/sextrans-2015-052090] [Medline: 27221475]
21. Delaney K, Branson B, Uniyal A, Kerndt P, Keenan P, Jafa K, et al. Performance of an oral fluid rapid HIV-1/2 test: experience from four CDC studies. AIDS 2006 Aug 01;20(12):1655-1660. [doi: 10.1097/01.aids.0000238412.75324.82] [Medline: 16868447]
22. Wesolowski LG, MacKellar DA, Facente SN, Dowling T, Ethridge SF, Zhu JH, et al. Post-marketing surveillance of OraQuick whole blood and oral fluid rapid HIV testing. AIDS 2006 Aug 01;20(12):1661-1666. [doi: 10.1097/01.aids.0000238413.13442.ed] [Medline: 16868448]
23. US Food and Drug Administration. 2012. OraQuick In-Home HIV Test: summary of safety and effectiveness URL: https://www.fda.gov/downloads/biologicsbloodvaccines/bloodbloodproducts/approvedproducts/premarketapprovalspmas/ucm312534.pdf [accessed 2017-04-21] [WebCite Cache ID 6ptNViHTD]
24. Katz D. Potential harms of HIV self-testing. 2015 Feb 23 Presented at: Conference on Retroviruses and Opportunistic Infections; February 23, 2015; Seattle, WA.
25. Katz D, Cassels S, Stekler J. Replacing clinic-based tests with home-use tests may increase HIV prevalence among Seattle men who have sex with men: evidence from a mathematical model. Sex Transm Dis 2014 Jan;41(1):2-9 [FREE Full text] [doi: 10.1097/OLQ.0000000000000046] [Medline: 24335742]
26. Campbell S, Klein R. Home testing to detect human immunodeficiency virus: boon or bane? J Clin Microbiol 2006 Oct;44(10):3473-3476 [FREE Full text] [doi: 10.1128/JCM.01511-06] [Medline: 17021069]
27. Weber K. OraQuick in-home HIV test. 2015 Feb 23 Presented at: Conference on Retroviruses and Opportunistic Infections; February 23, 2015; Seattle, WA.
28. Chan P, Towey C, Poceta J, Rose J, Bertrand T, Kantor R, et al. Online hookup sites for meeting sexual partners among men who have sex with men in Rhode Island, 2013: a call for public health action. Public Health Rep 2016;131(2):264-271 [FREE Full text] [doi: 10.1177/003335491613100210] [Medline: 26957661]
29. kontakt. New York, NY: kontakt.io, Inc; 2016 Jun 15. Extensive guide to Bluetooth low energy (BLE) beacons URL: https://kontakt.io/blog/extensive-guide-to-bluetooth-beacons/ [accessed 2017-04-21] [WebCite Cache ID 6ptOimAil]
30. Barnum C. Usability Testing Essentials: Ready, Set...Test!. Burlington, MA: Morgan Kaufmann; Nov 5, 2010.
31. Yen P, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc 2012;19(3):413-422 [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]

32. Bavinton BR, Brown G, Hurley M, Bradley J, Keen P, Conway DP, et al. Which gay men would increase their frequency of HIV testing with home self-testing? AIDS Behav 2013 Jul;17(6):2084-2092. [doi: 10.1007/s10461-013-0450-z] [Medline: 23525790]
33. Aidala A, Cross JE, Stall R, Harre D, Sumartojo E. Housing status and HIV risk behaviors: implications for prevention and policy. AIDS Behav 2005 Sep;9(3):251-265. [doi: 10.1007/s10461-005-9000-7] [Medline: 16088369]
34. Beijer U, Wolf A, Fazel S. Prevalence of tuberculosis, hepatitis C virus, and HIV in homeless people: a systematic review and meta-analysis. Lancet Infect Dis 2012;12(11):859-870. [doi: 10.1016/S1473-3099(12)70177-9]

Abbreviations

BLE: Bluetooth low energy
HIV: human immunodeficiency virus
HBST: home-based self-testing
IoT: Internet-of-Things
MSM: men who have sex with men

Edited by G Eysenbach; submitted 12.08.16; peer-reviewed by J Stekler, G Prestage; comments to author 26.11.16; revised version received 24.01.17; accepted 25.03.17; published 08.05.17.

Please cite as:
Wray T, Chan PA, Simpanen E, Operario D
eTEST: Developing a Smart Home HIV Testing Kit that Enables Active, Real-Time Follow-Up and Referral After Testing
JMIR Mhealth Uhealth 2017;5(5):e62
URL: http://mhealth.jmir.org/2017/5/e62/
doi:10.2196/mhealth.6491
PMID:28483744

Original Paper

Text Messaging and Mobile Phone Apps as Interventions to Improve Adherence in Adolescents With Chronic Health Conditions: A Systematic Review

Sherif M Badawy[1,2], MD, MS, MBBCh; Leonardo Barrera[3], MPH; Mohamad G Sinno[4], MD; Saara Kaviany[5], DO; Linda C O'Dwyer[6], MA, MSLIS; Lisa M Kuhns[7], MPH, PhD

[1]Ann & Robert H. Lurie Children's Hospital of Chicago, Department of Pediatrics, Division of Hematology, Oncology and Stem Cell Transplant, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

[2]Zagazig University Faculty of Medicine, Department of Pediatrics, Division of Hematology and Oncology, Zagazig, Egypt

[3]Ann & Robert H. Lurie Children's Hospital of Chicago, Department of Pediatrics, Division of Hematology, Oncology and Stem Cell Transplant, Chicago, IL, United States

[4]University of Kansas School of Medicine, Department of Pediatrics, Wichita, KS, United States

[5]Advocate Children's Hospital, Department of Pediatrics, Oak Lawn, IL, United States

[6]Northwestern University Feinberg School of Medicine, Galter Health Sciences Library, Chicago, IL, United States

[7]Ann & Robert H. Lurie Children's Hospital of Chicago, Department of Pediatrics, Division of Adolescent Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

Corresponding Author:
Sherif M Badawy, MD, MS, MBBCh
Ann & Robert H. Lurie Children's Hospital of Chicago
Department of Pediatrics, Division of Hematology, Oncology and Stem Cell Transplant
Northwestern University Feinberg School of Medicine
225 E Chicago Ave, Box 30
Chicago, IL, 60611
United States
Phone: 1 3122274836
Fax: 1 3122279376
Email: sbadawy@luriechildrens.org

Abstract

Background: The number of adolescents with chronic health conditions (CHCs) continues to increase. Medication nonadherence is a global challenge among adolescents across chronic conditions and is associated with poor health outcomes. While there has been growing interest in the use of mHealth technology to improve medication adherence among adolescents with CHCs, particularly text messaging and mobile phone apps, there has been no prior systematic review of their efficacy.

Objective: The purpose of this review was to systematically evaluate the most recent evidence for the efficacy of text messaging and mobile phone apps as interventions to promote medication adherence among adolescents with CHCs.

Methods: PubMed, Embase, CENTRAL, PsycINFO, Web of Science, Google Scholar, and additional databases were searched from 1995 until November 2015. An additional hand search of related themes in the Journal of Medical Internet Research was also conducted. The Preferred Reporting Results of Systematic Reviews and Meta-Analyses guidelines were followed. Two reviewers independently screened titles/abstracts, assessed full-text articles, extracted data from included articles, and assessed their quality using Grades of Recommendation, Assessment, Development, and Evaluation criteria. Included studies were described in original research articles that targeted adherence in adolescents with CHCs (12-24 years-old).

Results: Of the 1423 records examined, 15 met predefined criteria: text messaging (n=12) and mobile phone apps (n=3). Most studies were performed in the United States (11/15, 73%), were randomized-controlled trials (8/15, 53%), had a sample size <50 (11/15, 73%), and included adherence self-report and/or biomarkers (9/15, 60%). Only four studies were designed based on a theoretical framework. Approaches for text messaging and mobile phone app interventions varied across studies. Seven articles (7/15, 47%) reported significant improvement in adherence with moderate to large standardized mean differences. Most of the included studies were of low or moderate quality. Studies varied in sample size, methods of adherence assessment, and definition of adherence, which prohibited performing a meta-analysis.

Conclusions: The use of text messaging and mobile phone app interventions to improve medication adherence among adolescents with CHCs has shown promising feasibility and acceptability, and there is modest evidence to support the efficacy of these interventions. Further evaluation of short- and long-term efficacy and cost-effectiveness of these interventions is warranted given the early and evolving state of the science.

(JMIR Mhealth Uhealth 2017;5(5):e66) doi:10.2196/mhealth.7798

KEYWORDS

adolescents; medication adherence; chronic health conditions; chronic medical conditions; text messaging; mobile phone apps; smartphone apps; smartphone applications

Introduction

In the United States, about 15-20% of children and adolescents have chronic health conditions (CHCs) (eg, asthma, diabetes), a number that has doubled in the past 20 years accompanied by increased health care expenses [1,2]. The increased rate of children and adolescents with CHCs is mainly driven by the rising prevalence of asthma and obesity, and the advances in medical care that have led to improved survival over time [3,4]. Adolescents with CHCs have specific health needs and contend with daily challenges involving their illness-related routine, including administration of daily or weekly medications, diet restrictions, lifestyle changes, laboratory monitoring, and outpatient follow-up with medical teams [5].

Medication adherence in particular is an important component of the treatment regimen as it often drives long-term outcomes. It is also a global public health problem, and it represents a barrier to achieving optimal health as a primary cause of treatment failure and avoidable complications [5]. Medication nonadherence rates are estimated to be 50-75% among pediatric patients with CHCs, with some evidence of lower adherence among adolescents [5,6]. Medication nonadherence has been associated with poor health-related quality of life (HRQOL) scores, increased morbidity and mortality, and increased health care utilization with an estimated US $100-300 billion of annual avoidable health care costs [6-13]. Engaging adolescents with CHCs in self-management skill building, including medication adherence, has long-term benefits [14-16]. Although treatment regimen and monitoring requirements vary across pediatric CHCs, most adolescents with CHCs perceived adherence barriers as multifaceted, but with common attributes across conditions [17].

Earlier systematic reviews and meta-analyses of pediatric patients with CHCs have shown evidence of a positive impact of interventions on medication adherence, HRQOL, and family functioning as well as reduction in health care utilization [18-24], although with primarily small effect sizes and methodological limitations. There has been growing interest in the use of technology to improve medication adherence and self-management skills in the last few years. Adolescents have ubiquitous access to mobile technology, in particular text messaging and mobile phone apps, across levels of social position and status [25-27]. The adoption of these technologies by adolescents has opened up new opportunities to link patients with their providers and to improve self-management and medication adherence. A recent review examined the efficacy of mobile apps in improving self-management skills, not specifically medication adherence, among adolescents with a physical CHC or long-term condition. However, they were not able to draw concrete conclusions because of the lack of evidence-based studies and the heterogeneity of the included studies [28].

The purpose of this review is to systematically evaluate the most recent evidence for the efficacy of text messaging and mobile phone app interventions in promoting medication adherence among adolescents with CHCs. We focused on text messaging and mobile phone app interventions in particular because these technologies are the most widely and frequently used by adolescents and are thus most likely to serve as the basis for future intervention development.

Methods

We followed the guidelines for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) in the reporting of evidence across the studies we reviewed (Multimedia Appendix 1) [29]. This review was registered with the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42015025907) [30].

Article Retrieval

A librarian (LO) collaboratively developed the highly sensitive medical subject headings (MeSH) term based search strategies with other review authors (SB, LK) and from July to September 2015 ran searches in the following databases: PubMed MEDLINE; Embase; Cochrane Central Register of Controlled Trials (CENTRAL) on the Wiley platform; the Cumulative Index to Nursing and Allied Health Literature (CINAHL) (EBSCO); PsycINFO (EBSCO); Web of Science; Center for Review and Dissemination; and Inspec (EBSCO). Additional searches were run in November 2015 using the following sources: Proquest Dissertations; Scopus; ClinicalTrials.gov; World Health Organization Clinical Trials; Controlled-Trials.org; Institute of Electrical and Electronics Engineers Explore; and Google Scholar. Search strategies for all databases except MEDLINE were adapted from the PubMed MEDLINE strategy. All databases were searched back to 1995, which is a point in time when access to mobile phones began to increase rapidly. No language limits were applied. The search strategy specified keywords, including text messaging, phones, mobile apps, and portable software combined with adherence or compliance, and search terms related to child, pediatric, adolescents, and youth. We also reviewed the search strategies of previous studies to include additional terms. See Multimedia Appendix 2 for complete search strategies in each database. An

additional hand search of related themes in the *Journal of Medical Internet Research* was also conducted. We also attempted to identify additional studies by searching the reference lists of key studies and relevant systematic reviews. We contacted authors of included publications to obtain additional studies meeting the inclusion criteria.

Study Selection

The inclusion criteria were as follows: (1) adolescents (12-24 years old) [31] with a chronic illness requiring long-term daily or weekly medications for ≥12 months, (2) original research articles, (3) studies that were either randomized controlled trials, quasi-experimental studies, or pilot/feasibility studies (including single arm, pre-posttest), (4) text messaging or mobile phone based interventions (app or mobile intervention), and (5) medication adherence as the primary or secondary outcome. The exclusion criteria included (1) mean or median age of entire study cohort in the study was either <12 years old, >24 years old, or not specified, (2) adolescent participants were not the target of the intervention (eg, intervention targets babies born to adolescent mothers with human immunodeficiency virus or targets parents of adolescents), (3) text messaging and mobile phone apps as interventions focused on disease monitoring or ecological momentary assessment, but not meant to improve medication adherence, or (4) technology-based interventions other than text messaging and mobile phone apps, such as Web- or Internet-based interventions, personal digital assistant, etc.

Data Extraction

We used a standardized form for data extraction. Data items in the extraction form included the following: first author's name; publication year; country; condition or disease focus of the study; participants' age; study design; duration of intervention and follow-up; components of text messaging or mobile phone app interventions; control group (if applicable); adherence measures; adherence rates; other outcome measures such as disease-related outcomes of morbidity and mortality, HRQOL, and self-efficacy or self-management skills; and theoretical framework. Two authors coded all included articles individually, and then the lead author (SB) independently reviewed all codes. Disagreements were resolved by discussion or by consultation with a third author, if needed.

Quality Assessment and Strength of the Evidence

Studies described in each article were evaluated for the quality of evidence using the GRADE approach (Grades of Recommendation, Assessment, Development, and Evaluation) [32]. This method evaluates four different key domains including consistency, directness, risk for bias, and precision of the evidence. Two authors graded all included articles individually, and then the senior author (LK) independently reviewed all grades. Disagreements were resolved by discussion or by consultation with a third author, if needed.

Data Analysis

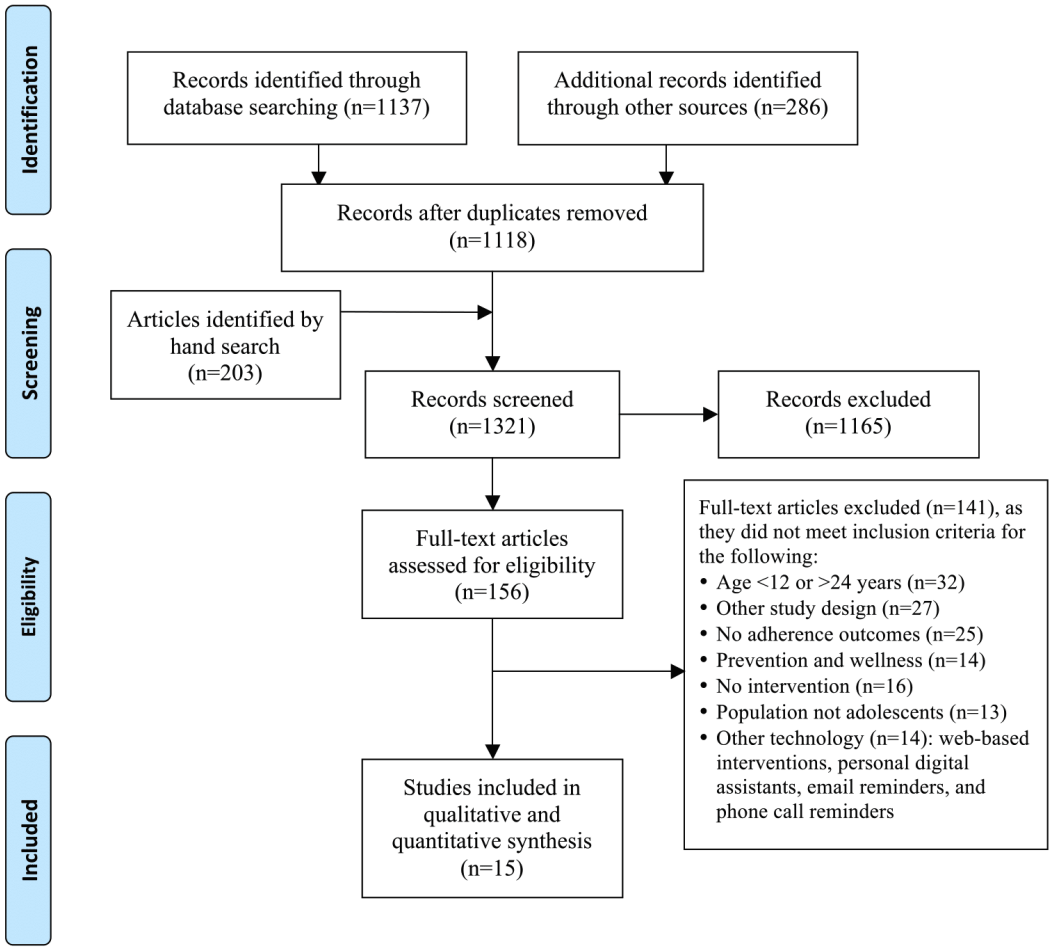
Data were analyzed quantitatively and qualitatively. Our primary outcome measure was mean change in medication adherence rate, and data were analyzed for those who had baseline and follow-up values. We also analyzed mean change in adherence-related laboratory markers. Standardized mean differences (SMD) with 95% confidence intervals were calculated—using means and standard deviations—to evaluate the efficacy of text messaging and mobile phone based interventions in improving subjective and objective measures of adherence, including adherence rates and adherence-related laboratory markers [33]. Data were analyzed using Stata 13.

Results

Literature Search

A total of 1423 citations were retrieved; 1137 in the July-August 2015 search and 286 in the search in November 2015. After removing duplicates, 1118 original articles remained (see Figure 1). Two authors (SB and LK) independently screened the article titles and abstracts of 1118 records against inclusion criteria, and a total of 156 records met all predefined inclusion criteria. Two authors (SB and LK) then independently reviewed the full text of these articles in detail against the exclusion criteria, and 141 articles were excluded. A total of 15 articles met all predefined criteria to be included in this review [34-48]. We did not identify any non-English articles that met our inclusion criteria. The study flowchart and reasons for exclusion of full text papers were documented in an adapted PRISMA study flowchart (see Figure 1) [49].

Figure 1. Flow of studies through the review according to PRISMA guidelines.



Study Characteristics

Although our search included studies published since 1995, no eligible studies were identified before 2005, with most studies (12/15, 80%) published since 2010 (Table 1) [34-39,41-44,46,48]. Most studies were performed in the United States (in whole or in part; 11/15, 73%) [34,36-38,41,42,44-48]. Studies targeted eight different CHCs, including those of adolescents with acne (n=2) [34,39], asthma (n=1) [47], depression (n=1) [42], diabetes mellitus (n=4) [35,40,43,46], human immunodeficiency virus (n=2) [37,41], liver transplant (n=2) [44,45], sickle cell disease (n=2) [36,38], and systematic lupus erythematosus (n=1) [48]. Most studies were small in size (ie, 50 participants; 11/15, 73%) [34-37,42-48], and just over half (8/15, 53%) included samples of young adults (with a mean or median age ≥18 and ≤24 years old) [34,37,39,41-43,47,48]. In terms of study design, more than half were RCTs (8/15, 53%) [34,39-43,47,48], and the remainder were primarily single-arm nonrandomized trials (6/15, 40%) [35-37,44-46] or retrospective chart reviews (1/15, 7%) [38]. The duration of the studies varied: 2-4 weeks (2/15, 13%) [42,43], 12-16 weeks (5/15, 33%) [34,35,39,46,47], 24 weeks (3/15, 20%) [36,37,41], and 12 months or more (5/15, 33%) [38,40,44,45,48]. Only one study evaluated the durability of intervention effects in a crossover design with 6-month follow-up in one of the study arms after the intervention was discontinued [41]. Measures of medication adherence included self-report (9/15, 60%) or biomarkers (9/15, 60%), as well as other forms of monitoring (7/15, 47%), such as Medication Event Monitoring System caps, directly observed therapy, pill count, and pharmacy records abstraction. Three studies (20%) measured additional adherence behaviors, including completion of laboratory visits [44], clinic visits [48], and monitoring of peak expiratory flow values in patients with asthma [47]. Only four studies (27%) incorporated explicit theoretical approaches or frameworks into the model of intervention effects, including Theory of Planned Behavior [43], Gamification theory [35], and Social Cognitive Theory [40,41]. Most studies (12/15, 80%) were rated “low” or “moderate” according to the GRADE criteria (Table 1) [32], with low ratings primarily due to limitations in design as well as imprecision of results.

Table 1. Summary of included studies that focused on improving adherence in adolescents with CHCs using text message or mobile phone app interventions.

Source (condition)	Intervention (study design)	Age (years)	Participants	Adherence measure	Grade
Boker et al 2012, USA (acne) [34]	Text messages (RCT)	Mean (SD) (range): Entire cohort 22.7 (5.7) (12-35) Text 22.8 (5.6) (14-35) Control 22.5 (5.9) (12-32)	Total N: Enrolled (40) Lost follow-up (7) Final N=33: Intervention (15); Control (18)	Medication event monitoring system	Low
Fabbrocini et al 2014, Italy (acne) [39]	Text messages (RCT)	Mean age: Text 19.5 Control 18.5 Entire cohort range (14-28)	Total N: Enrolled (160) Lost follow-up (15) Final N=145: Intervention (74); Control (71)	7-day recall self-report	Low
Ostojic et al 2005, Croatia and USA (asthma) [47]	Text messages and in-person sessions (RCT)	Mean (SD): 24.6 (6.5)	Final N=16: Intervention (8); Control (8)	Self-report of daily medication use in a paper diary	Low
Hammonds et al 2015, USA (depression) [42]	Mobile phone app (RCT)	Mean (SD): 20.6 (4.3)	Total N: Enrolled (57) Lost follow up (10) Final N=40: Intervention (20); Control (20)	Pill count	Low
Cafazzo et al 2012, Canada (diabetes mellitus) [35]	Mobile phone app "*bant*" (pilot trial)	Mean (SD): 15.1 (1.3)	Total N=20: Withdrawal (2) Incomplete data (6) Final N=12, Intervention	Self-report Laboratory marker (glycosylated hemoglobin "HbA1c")	Moderate
Franklin et al 2006, UK (diabetes mellitus) [40]	Text messages "Sweet Talk" (ST) (RCT)	Age (median, range): CIT 12.7 (10.5-14.8) Conventional insulin therapy /ST 14.1 (11.7-15.6) Intensive insulin therapy/ST 12.6 (11.2-15.4)	Total N: Enrolled (92) No allocation (3) Lost follow-up (1) Discontinued text (13) Final N=90: Conventional insulin therapy (27) Conventional insulin therapy and ST (32) Intensive insulin therapy and ST (31)	Self-report Laboratory marker (HbA1c)	Moderate
Louch et al 2013, UK (diabetes mellitus) [43]	Text messages (RCT)	Range for entire cohort is 18-30	Total N: Enrolled (19) Lost follow-up (1) Final N=18: Intervention (8); Control (10)	Self-report	Low
Mulvaney et al 2012, USA (diabetes mellitus) [46]	Text messages "SuperEgo" (pilot trial with a historical control)	Mean (SD): Intervention 15.9 (2.9); Controls 15.8 (2.7)	Enrolled (28) Incomplete data (5) Final N=46: Intervention (23); Historically matched controls (23)	Laboratory marker (HbA1c)	Low
Dowshen et al 2012, USA (human immunodeficiency virus "HIV") [37]	Text messages (pilot trial, pre-post design)	Mean (range): 23 (14-29)	Total N: Enrolled (25) Lost follow-up (3) Technical issue (1) Final N=21	Self-report (visual analogue scale and AIDS Clinical Trials Group) Laboratory markers (viral load and CD4 cell count)	Low
Garofalo et al 2015, USA (HIV) [41]	Text messages (RCT)	Mean (SD): 24.1 (2.9) Median (range): 23 (18-29)	Final N=105: Control (51) Intervention (54)	Self-report (visual analogue scale) Laboratory markers (viral load and CD4 cell count)	Moderate

Source (condition)	Intervention (study design)	Age (years)	Participants	Adherence measure	Grade
McKenzie et al 2015, USA (liver transplant) [44]	Text messages (pilot trial with a historical control)	Median (range): 16 (12-20)	N=42	Laboratory testing participation rate	Low
Miloh et al 2009, USA (liver transplant) [45]	Text messages (pilot trial, pre-post design)	Median (range): 15 (1-27)	Total N: Enrolled (41) Dropout (17) Final N=24	Laboratory markers (tacrommilus levels and SD values)	Low
Creary et al 2014, USA (sickle cell disease) [36]	Text messages and Mobile Direct Observed Therapy "m-DOT" (pilot trial, pre-post design)	Mean (SD): 13.7 (6.3)	Total N=15 Final N=14	Observed adherence Self-report (Morisky Medication Adherence Scale) Medication possession ratio Laboratory markers (fetal hemoglobin and mean corpuscular volume)	Very low
Estepp et al 2014, USA (sickle cell disease) [38]	Text messages "SIMON" (retrospective study)	Median (range): 13.9 (12.1-16.1)	Total N=83 Final N=55	Medication possession ratio Laboratory markers (fetal hemoglobin and mean corpuscular volume)	Low
Ting et al 2011, USA (systemic lupus erythematosus) [48]	Text messages (RCT)	Mean (SD): 18.6 (2.5)	Final N=41: Intervention (19); Control (22)	Self-report (Medication Adherence Self-Report Inventory) Medication possession ratio Laboratory markers (hydroxychloroquine levels)	Low

Intervention Type

The majority of interventions used text messaging to promote medication adherence via reminders or motivational messages (12/15, 80%) [34,37-41,43-46,48]. Of these 12 studies, only one combined text messages with other in-person intervention components (educational sessions) [47]. Additional interventions used other mobile phone based approaches or apps (3/15, 20%), including multifunction mobile phone apps [35,42] and mobile phone based directly observed therapy [36].

Intervention Characteristics

Text Messaging Interventions

Text messaging interventions varied by frequency of messaging, message content, and added functionality. Most of text messaging interventions included one [36-38,40-43,45], two [34,39,48], or three [47,48] daily text reminders, or even more frequently in relation to meals [46]. Other studies provided reminders at different frequencies including weekly [40,41,46,47]; monthly, bi-monthly, or quarterly for laboratory monitoring [44]; and 1, 3, and 7 days before scheduled clinic appointments [48]. Most studies sent text message reminders that were customized to the patient's medication regimen or personal preferences in terms of both scheduling (ie, time/day) and message content [34,37-41,43-48]. In addition, four studies had patients and/or parents create the content of the text messages themselves [38,40,45,46], and two of these studies included the use of a text message pool or bank [40,46]. The most sophisticated intervention designs also included messages related to dosing, side effects, adherence barriers, goal setting, positive reinforcement [39,40,43,46], or targeted to theoretical constructs [43]. Additional functionalities included patient reporting of physiological information via text (eg, peak expiratory flow in asthma patients) [47]; prompting of text-back responses from patients (two-way text messaging approach) [34,37,41,44,45]; sending text messages to parents if patients did not respond to scheduled reminders [45]; and the ability to request messages from individuals within their social network via the intervention platform [35,46] or from a motivational support network [40]. None of the text message intervention studies included a reward system or scheduled incentives for participants. Table 2 describes the approach and the components of different text messaging interventions.

Table 2. Description of text message interventions.

Author/year (condition)	Intervention purpose	Intervention description
Boker et al, 2012 (acne) [34]	Improve adherence to recommended use of topical acne medication (text messages)	Text messages twice daily (Duac in AM, Gifferin in PM) for 12 weeks
		Customized electronic reminder schedule at a specific time based on patient preferences and anticipated time of each medication use
		2-way communication: patients asked to text back a reply if and when each application was completed
		Identical text with general content to all patients, varied only by starting with patient's first name
		Texts were sent through: www.LetterMeLater.com
Fabbrocini et al, 2014 (acne) [39]	Improve adherence to acne medications (text messages)	Texts twice daily for 12 weeks (11 consecutive days)
		Texts focused on frequently asked questions about acne medications, such as administration, daily dose, and side effects
		Texts were identical to all patients (no customization) covering 11 frequently asked questions
		Texts re-sent in same sequence every 11 days until end of 12 weeks
Ostojic et al, 2005 (asthma) [47]	Improve adherence to inhaled medications and peak expiratory flow (PEF) monitoring (text messages and in-person sessions)	Patients sent their PEF results daily via text messages for 16 weeks
		Data connected to a computer with software that automatically computed maximal, minimal, and mean PEF, PEF variability, and compliance
		PEF measurements 3 times daily with medication use and symptoms in paper diary
		Therapy was adjusted weekly by an asthma specialist according to peak expiratory flow meter (PEFM) values received daily from the patients
		1-hour asthma education session with specialist at clinic: discussed symptoms, asthma symptom score, indicators of control, medication use, and correct use of metered dose inhaler and PEFM
Louch et al, 2013 (diabetes mellitus) [43]	Improve insulin administration in young adults with type 1 diabetes; test moderation of intervention effect by personality factors (text messages)	Text messages sent daily (1-way communication) at 10 am for 14 days
		No customization of message content
		Text content was related to the correct insulin administration
		Text targeted constructs of the Theory of Planned Behavior: attitudes, subjective norms, perceived behavioral control, and intention

Author/year (condition)	Intervention purpose	Intervention description
Mulvaney et al, 2012 (diabetes mellitus) [46]	Motivate patients and remind them with their diabetes self-care tasks (text messages "SuperEgo")	8-12 text messages/week for 12 weeks
		Scheduled just before and after mealtimes and before bedtime
		Customization with users able to alter timing and frequency of messages through a website
		Messages could be scheduled in a 1-way communication at specific times of day within 15 minute increments and automated to be sent once only, or repeated based on participant preferences, such as daily, weekly, or on weekends
		Individually tailored messages: 75% of messages tailored to the top 3 patient-specific adherence barriers reported; and 25% of messages were randomly selected from the remaining message pool
		Four functions were included in the system: assessment, message selection, message scheduling, and requests for messages from others
		Text messages were created in collaboration with 96 adolescents with diabetes mellitus and no messages were repeated
		Participants could add their own messages, search for messages, and delete, change, or reschedule them using their mobile phone
		Participants could also search for and select messages that were associated with a particular goal
		Participants could ask other SuperEgo users for messages relating to a specific goal and then schedule that message for themselves
		Messages could be specified as private or public
		Participants could nominate people as part of their support team by entering that person's email address into the system to contribute messages for patients
Franklin et al, 2006 (diabetes mellitus) [40]	Improve patients' self-efficacy and glycemic control, and enhance their uptake of intensive insulin therapy (text messages "Sweet Talk")	Text message reminders for 12 months
		Weekly reminders of the goal set in clinic, and daily reminders providing tips, information, or reminders to reinforce this goal
		Text messages automated, scheduled, and designed to offer regular support and optimize their self-management and control
		Database of over 400 messages that encompass the four main diabetes self-management tasks (insulin injections, blood-glucose testing, healthy eating, and exercise)
		Messages tailored based on patients goals and patients' age, sex, and diabetes regimen
		Occasional text "newsletters" regarding topical diabetes issues
		Motivational support network
Dowshen et al, 2012 (HIV) [37]	Improve adherence to antiretroviral therapy among youth (text messages)	Daily 2-way text messages for 24 weeks
		Delivered at prespecified time schedule
		Personalized content, patients were encouraged to develop messages that maintain their confidentiality
		Interactive with follow-up messages with patients responding with number (1) if they took their medication and (2) if they did not
		Participants could reach out to study coordinator at any time to change the message or to reprogram the message if their mobile service was interrupted
		Texts were sent through: http://www.intelecare.com/

Author/year (condition)	Intervention purpose	Intervention description
Garofalo et al, 2015 (HIV) [41]	Improve adherence to antiretroviral therapy among poorly adherent youth (text messages)	Daily text messages reminder for 6 months
		Initial messages were followed by a second message 15 min later to check if patients took their medications in a 2-way communication
		Personalized by subject for both content and schedule to be timed with medication doses
		Customization with initial message and follow-up messages were designed by the youth themselves
		Texts content were culture and identity sensitive and meaningful to participants
		Texts content used more indirect language to maintain confidentiality
		Motivational or encouraging follow-up messages were randomly sent to participants based on their affirmative or negative response
		Participants were encouraged to delete messages after taking medication and to use passcode protection to maintain phone confidentiality
		Sent/received and failed/invalid messages were summarized in weekly reports and sent to research staff to follow up with participants
		Texts were sent through: http://www.remedyhealthmedia.com/
Miloh et al, 2009 (liver transplant) [45]	Improve adherence to immunosuppressant medications (text messages)	Mean duration of the study 13 (SD 1.5) months
		Text schedule was customized at day/time specified by user
		2-way communication where patients were expected to respond to text message to confirm medication intake; if no response within 15 min to 1-hour caregiver notified via text
		Text messages were sent to the person responsible for medication intake (patient or caregiver)
		Patients or their caregivers registered and entered their information into texting platform (MediM system) with a personal password
		Entered information included patient's name and mobile phone number, caregiver's name/ nickname and mobile phone number, the medication name and frequency, and the exact times of text messages they want to receive
		Participants did not enter medication dose as that might change based laboratory test results
		No customization as text messages content was the same for all patients
		Text messages were read: "Take [name of medication] at [set time]. To confirm intake, press REPLY, type CARE 1, and press SEND."
		Participants reimbursed for all text messages costs during the study
		Texts were sent through: http://www.carespeak.com/corp/
McKenzie et al, 2015 (liver transplant) [44]	Improve participation in laboratory testing among youth (text messages)	Automated laboratory tests text message reminders for 12 months
		Text message timed with lab tests (monthly, bimonthly, quarterly) as reminder to complete lab tests
		Text message reminders sent first Monday of each month when testing was due
		On last Monday of the month, patients received a message about laboratory testing completion
		Same message content for all patients
		2-way communication as patients replied back as yes/no responses
		No reimbursement for the cost of text messages, but all participants had unlimited text plans
		Mobile phone numbers entered into a secure website with secure-password
		$31/month to maintain the intervention or website domain

Author/year (condition)	Intervention purpose	Intervention description
Estepp et al, 2014 (sickle cell disease) [38]	Improve adherence to hydroxyurea therapy (text messages: SIMON)	Scheduled daily text message reminders for 12 months
		Customizable for content, frequency, and duration
		Participants created their own messages
		Changes in text messages regimen checked every 3-4 clinic visits
		Messages delivery was monitored (received and undelivered) and participants could optionally reply
		Messages sent through a Web-based app
Ting et al, 2011 (systemic lupus erythematosus) [48]	A. Visit adherence intervention Improve adherence to scheduled clinic visits (text messages)	Text reminders sent 7, 3, and 1 day prior to each scheduled f/u clinic appointment
		Mean duration of the study 12 (SD 5) months
		Content was individualized for each patient and included the scheduled appointment time
		If a patient didn't schedule follow-up appointment within 2–3 weeks after completed clinic visit, they would get a text reminder to do so
	B. HCQ adherence intervention Improve adherence to use of hydroxychloroquine (text messages)	Standardized daily text reminders for hydroxychloroquine intake daily or twice daily
		Mean duration of the study 12 (SD 5) months
		Text reminders at set time of day, according to hydroxychloroquine schedule
		Printed information sheet that was given to the standard of care group

Mobile Phone App Interventions

Mobile phone based interventions included native apps for delivery of medication reminders [42]; a multifunction app that includes integration of a wireless device for physiological measurement and visualization (ie, glucometer); self-monitoring alerts and prompts for gamification features to incentivize engagement with the app, with a secure network for peer communication [35]; and a multifunction app that includes daily alert messages, creation and sharing of patient videos to directly observe adherence to therapy with feedback/follow-up, positive feedback/reinforcement messages, and incentives for adherence [36]. Table 3 describes the approach and the components of different mobile phone app interventions.

Study Outcomes

Of the 15 studies reviewed, 7 (47%) demonstrated statistically significant differences in medication adherence or related health outcomes [36,37,41,43-45,47]. The majority of the studies included in this review provided enough information to calculate standardized outcomes, such as SMDs. Most studies reported overall moderate to large SMDs of subjective and objective markers of adherence; however, most SMDs had wide 95% confidence intervals (Table 4). Several studies combined data related to the assessment of the efficacy or the usability/feasibility of different interventions reporting high levels of satisfaction and few technical or feasibility problems [35-37,40,41,44-46]. Additional results of each individual study are summarized in Multimedia Appendix 3 for text message interventions, and Multimedia Appendix 4 for mobile phone app interventions.

Table 3. Description of mobile phone app interventions.

Author/year (condition)	Intervention purpose	Intervention description
Hammonds et al, 2015 (depression) [42]	Improve adherence to antidepressant medications among college students using mobile phone app reminders	Medication reminders through mobile phone app for 4 weeks
		Entered prescribed information for medication doses
		Patients indicate when they had taken their medication by responding to reminders received
Cafazzo et al, 2012 (diabetes mellitus) [35]	Improve self-management among youth (mobile phone app " *bant* ")	App exposure for 12-week pilot study
		Adapter that allows a OneTouch UltraMini glucometer to communicate via Bluetooth, allowing the transfer of blood glucose reading wirelessly, to the iPhone device running the mobile phone app, " *bant* "
		Analysis tools assess the data soon after transfer to give adolescents real-time feedback
		Data analysis and trending screens display the percentage of blood glucose levels that are in range at specific times
		Communication with peers in an app-secure community area as a support network
		Rewards algorithm with point allocations
Creary et al, 2014 (sickle cell disease) [36]	Improve adherence to hydroxyurea (m-DOT, multidimensional strategy for 6 months)	Alert reminders: automated daily alerts to remind patients to take hydroxyurea; alert sent at time preferred by patients; alerts stopped when a video is submitted; up to 4 text messages and email were sent daily
		Videos: participants created daily videos of them taking hydroxyurea; videos submitted electronically to the secure study website; captured by mobile phones or computers; included participants' study ID; self-recorded videos for children 12 years or older, younger patients had assistance from parents
		Feedback: submitted videos were reviewed by research team within 72 hours of receipt; text and email feedback was sent to participants if they missed ≥2 video submissions in a 30-day period; participants were called if they missed ≥3 video submissions in a 30-day period; positive reinforcement (≥2 text messages or emails) was provided if they participants had adherence of ≥90%
		Incentives: if participants achieved ≥90% of adherence to hydroxyurea for each 30-day period, they would receive $1/day

Discussion

Principal Findings

Medication nonadherence is a widespread problem in pediatric CHCs, and among adolescents in particular. In this systematic review, we identified 15 studies that met all our inclusion criteria. Our results suggest that there is modest evidence to support the efficacy of text messaging and mobile phone apps as interventions to improve medication adherence in adolescents with CHCs. Most of the included studies were of low to moderate quality because of methodological limitations, imprecision of results, or both. The included studies showed evidence of the acceptability and feasibility of text messaging and mobile phone app interventions, suggesting a potentially promising area of intervention development and further investigations in the near future to better understand their efficacy and cost-effectiveness.

Our findings suggest moderate SMDs for most included interventions, which is consistent with earlier reports of adherence-enhancing interventions (ie, non-technology specific). However, given the heterogeneity of the included studies, the observed moderate effect size should be interpreted with caution [18,23]. In contrast, Pai and McGrady reported small effect sizes in a systematic review of adherence-promoting interventions [20]. In our review, the quality of the included studies was low to moderate, which was similar to a recent review of findings for technology-mediated interventions for treatment adherence across all age groups [19], and more specifically for adolescents with chronic physical conditions [28].

Table 4. Effect size *d* for the main outcomes of included studies.

Source (intervention)	Study adherence outcomes	*d* (95% CI)[a]
Boker et al, 2012 (text messaging) [34]	Medication event monitoring system	No data available
Fabbrocini et al, 2014 (text messaging) [39]	7-day recall self-report	No data available
Ostojic et al, 2005 (text messaging and in-person sessions) [47]	Self-report of daily inhaled corticosteroids	0.35 (-0.64 to 1.4)
	Self-report of daily beta2-agonist	0.7 (-0.31 to 1.71)
Hammonds et al, 2015 (mobile phone app) [42]	Pill count of antidepressants	0.31 (-0.21 to 0.83)
Cafazzo et al, 2012 (mobile phone app) [35]	Self-report using personal blood glucose meters	0.11 (-0.69 to 0.91)
	Laboratory markers using glycosylated hemoglobin	0.45 (-0.36 to 1.26)
Franklin et al, 2006 (text messaging) [40]	Self-report using a visual analogue scale	0.38 (-0.14 to 0.89)
	Laboratory marker using glycosylated hemoglobin	0.12 (-0.4 to 0.63)
Louch et al, 2013 (text messaging) [43]	Self-report of insulin administration	1.1 (0.11 to 2.1)[b]
Mulvaney et al, 2012 (text messaging)	Laboratory markers using glycosylated hemoglobin	0.5 (-0.1 to 1.1)
Dowshen et al, 2012 (text messaging) [37]	Self-report using visual analogue scale	1.43 (0.75 to 2.11)[b]
	Self-report using AIDS Clinical Trials Group questionnaire	0.86 (0.22 to 1.49)[b]
	Laboratory marker using viral load	0.43 (-0.18 to 1.04)
	Laboratory marker using CD4 cell count	0.19 (-0.42 to 0.79)
Garofalo et al, 2012 (text messaging) [41]	Self-report using visual analogue scale	0.49 (0.11 to 0.89)[b]
	Laboratory markers using viral load	0.19 (-0.18 to 0.58)
McKenzie et al, 2015 (text messaging) [44]	Laboratory testing participation rate	0.66 (0.22 to 1.1)[b]
Miloh et al, 2009 (text messaging) [45]	Laboratory markers using tacromilus – overall	1.23 (0.62 to 1.85)[b]
	Laboratory markers using tacromilus – one medication	1.02 (0.42 to 1.6)[b]
	Laboratory markers using tacromilus – two medications	1.39 (0.77 to 2.03)[b]
	Laboratory markers using tacromilus – three medications	2.11 (1.41 to 2.82)[b]
Creary et al, 2014 (text messaging and m-DOT) [36]	Medication possession ratio	1.04 (0.25 to 1.83)[b]
	Laboratory markers using fetal hemoglobin	0.1 (-0.31 to 0.51)
	Laboratory markers using mean corpuscular volume	0.54 (-0.22 to 1.29)
Estepp et al, 2014 (text messaging) [38]	Medication possession ratio	0.07 (-0.31 to 0.44)
	Laboratory markers using fetal hemoglobin	0.1 (-0.31 to 0.51)
	Laboratory markers using mean corpuscular volume	0.18 (-0.24 to 0.59)
Ting et al, 2011 (text messaging) [48]	Self-report using medication adherence inventory	No data available
	Medication possession ratio	No data available
	Laboratory marker for hydroxychloroquine	No data available

[a]Positive effect size value means improvement in a study outcome, while a negative one means worsening outcome.

[b]Statistically significant $P<.05$.

While text messaging and mobile phone app interventions offer a straightforward approach to address adherence to medications among adolescents, with broad application across CHCs, some challenges exist. None of the included studies measured the long-term durability of intervention effects across randomized conditions; thus, there is a need to establish the exposure or dosage needed to sustain long-term effects. In addition, the characteristics of the included studies also suggest that there is a need for improvement in intervention design; only four included theoretical models or approaches to target specific mechanisms of action to optimize efficacy [35,40,41,43]. McGrady and colleagues have recommended that articulation of the underlying mechanism of action for adolescent-specific adherence interventions is an advancement much needed to bring developmental and behavioral specificity to this growing area of research [50]. Furthermore, only two studies measured potential moderators of the intervention effect [41,43].

Several of the studies measured feasibility and acceptability outcomes and found high levels of satisfaction and few feasibility issues, which are promising for advancements in these technologies and consistent with earlier reports [19,21,28]. The high level of satisfaction is noteworthy given the frequency of messaging, which for most studies included at least daily messages, reflecting a relatively high tolerance in this group for intervention. The evidence of feasibility in these studies suggests that adolescent-specific text messaging and mobile phone based approaches may be an important and promising area of future intervention development. Given methodological limitations in the studies reviewed, larger studies with long-term outcomes are warranted, particularly with sufficient power for clinically important outcomes. Recent evidence supports the efficacy of text message and mobile phone app interventions to promote preventive behavior among adolescents [51]. Furthermore, in addition to efficacy data, cost-effectiveness is another important aspect of intervention evaluation [9,22,52]. The cost to develop and maintain each intervention can be a barrier to widespread use of these interventions. Cost also impacts the variability in patients' access to technologies. Formal economic evaluation of different interventions will inform health care authorities to decide whether adoption of such interventions would be economically efficient [22,52]. In a related systematic review, we found insufficient evidence regarding the cost-effectiveness of text messaging and mobile phone apps to promote adherence in adolescents with CHCs [52]. Our findings highlight the need for further investigation of cost-effectiveness to inform the scalability, sustainability, and future cost savings of such approaches [52].

Strengths and Limitations

Our systematic review has a number of strengths. First, we conducted our review following the recommendations for rigorous systematic reviews methodology [32,49,53]. Second, we used a highly sensitive search strategy guided by a librarian information specialist and had no language restrictions in order to minimize publication bias by identifying as many relevant studies as possible. Additional resources were also searched including published systematic reviews, clinical trial registries, and different electronic databases. Third, although we limited our search to studies published since 1995, there were no eligible studies identified before 2005. Therefore, the possibility that we have missed earlier studies is very small. Finally, 2 authors completed the review process independently at all stages.

Our systematic review of the literature should be considered within the context of some potential methodological limitations. First, similar to any other systematic literature review, although our search criteria were designed to be comprehensive, it is possible that we missed relevant articles. Second, we included only articles published in peer-reviewed journals, and publication bias with the tendency to report positive study results cannot be excluded [54]. Third, a number of studies reported insufficient information, the definition of adherence varied, and the study sample size and ages as well as methods of adherence assessment used in the included studies were heterogeneous. These limitations prohibited a meta-analysis from being performed [55]. Finally, many of the included studies had relatively small sample sizes.

Conclusions

The number of adolescents with CHCs continues to increase and medication nonadherence is a clear challenge. The use of text messaging and mobile phone app interventions to improve medication adherence among adolescents with CHCs has shown promising feasibility and acceptability; however, there is modest evidence to support their efficacy. Further evaluation of short- and long-term efficacy and cost-effectiveness of these interventions is warranted. In addition, better understanding of barriers to medication adherence would inform further development of text message and/or mobile phone app interventions to improve adherence and health outcomes in adolescents with CHCs, such as sickle cell disease [56,57]. Adolescents are frequent users of text messaging and mobile phone apps, and engaging adolescents with CHCs in their self-management is critical for improving long-term outcomes. Seeking adolescents' perspectives could enhance uptake and long-term engagement, while minimizing patient fatigue. The currently available data from medication adherence studies using text messaging and mobile phone app interventions are heterogeneous, particularly in relation to process and outcomes measures, which limit the evidence generated and conclusions that can be drawn from those studies. Nevertheless, consistent use of Web-based and mobile health interventions reporting guidelines [58] would enhance comparative research across studies [59]. The functionalities of different mobile technologies continue to improve with gradually decreasing cost suggesting potential economies of scale where interventions could be delivered to large populations.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA checklist.

[PDF File (Adobe PDF File), 67KB - mhealth_v5i5e66_app1.pdf]

Multimedia Appendix 2

Search strategies.

[PDF File (Adobe PDF File), 151KB - mhealth_v5i5e66_app2.pdf]

Multimedia Appendix 3

Text message intervention outcomes.

[PDF File (Adobe PDF File), 58KB - mhealth_v5i5e66_app3.pdf]

Multimedia Appendix 4

Mobile phone app intervention outcomes.

[PDF File (Adobe PDF File), 32KB - mhealth_v5i5e66_app4.pdf]

References

1. Van CJ, Gortmaker SL, Perrin JM. Dynamics of obesity and chronic health conditions among children and youth. JAMA 2010 Feb 17;303(7):623-630. [doi: 10.1001/jama.2010.104] [Medline: 20159870]
2. Berry JG, Hall M, Hall DE, Kuo DZ, Cohen E, Agrawal R, et al. Inpatient growth and resource use in 28 children's hospitals: a longitudinal, multi-institutional study. JAMA Pediatr 2013 Feb;167(2):170-177 [FREE Full text] [doi: 10.1001/jamapediatrics.2013.432] [Medline: 23266509]
3. Sawyer SM, Afifi RA, Bearinger LH, Blakemore S, Dick B, Ezeh AC, et al. Adolescence: a foundation for future health. Lancet 2012 Apr 28;379(9826):1630-1640. [doi: 10.1016/S0140-6736(12)60072-5] [Medline: 22538178]
4. Smith JM, Martz K, Blydt-Hansen TD. Pediatric kidney transplant practice patterns and outcome benchmarks, 1987-2010: a report of the North American Pediatric Renal Trials and Collaborative Studies. Pediatr Transplant 2013 Mar;17(2):149-157. [doi: 10.1111/petr.12034] [Medline: 23281637]
5. Sabate E. Adherence to Long-term Therapies: Evidence for Action. 2003. World Health Organization URL: http://apps.who.int/iris/bitstream/10665/42682/1/9241545992.pdf [accessed 2016-10-15] [WebCite Cache ID 6pIt6ezF8]
6. Rapoff M. Adherence to pediatric medical regimens. In: Issues in Clinical Child Psychology. 2nd ed. New York: Springer Science+Business Media; 2010.
7. DiMatteo MR. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Med Care 2004 Mar;42(3):200-209. [Medline: 15076819]
8. DiMatteo MR, Giordani PJ, Lepper HS, Croghan TW. Patient adherence and medical treatment outcomes: a meta-analysis. Med Care 2002 Sep;40(9):794-811. [doi: 10.1097/01.MLR.0000024612.61915.2D] [Medline: 12218770]
9. McGrady ME, Hommel KA. Medication adherence and health care utilization in pediatric chronic illness: a systematic review. Pediatrics 2013 Oct;132(4):730-740 [FREE Full text] [doi: 10.1542/peds.2013-1451] [Medline: 23999953]
10. Davis K. Agency for Healthcare Research and Quality. 2011. Statistical brief #343: Health care expenses and utilization for children with special health care needs, 2008. Estimates for the U.S. civilian noninstitutionalized population URL: https://meps.ahrq.gov/data_files/publications/st343/stat343.shtml [accessed 2017-03-28] [WebCite Cache ID 6pIuNu8US]
11. Fredericks EM, Magee JC, Opipari-Arrigan L, Shieck V, Well A, Lopez MJ. Adherence and health-related quality of life in adolescent liver transplant recipients. Pediatr Transplant 2008 May;12(3):289-299. [doi: 10.1111/j.1399-3046.2008.00901.x] [Medline: 18282211]
12. New England Health Institute. NEHI research shows patient medication nonadherence costs health care system $290 billion annually. 2009. URL: http://www.nehi.net/news/344-nehi-research-shows-patient-medication-nonadherence-costs-health-care-system-290-billion-annually/view [accessed 2017-03-27] [WebCite Cache ID 6pIv2YOqv]
13. Sokol MC, McGuigan KA, Verbrugge RR, Epstein RS. Impact of medication adherence on hospitalization risk and healthcare cost. Med Care 2005 Jun;43(6):521-530. [Medline: 15908846]
14. Staples B, Bravender T. Drug compliance in adolescents: assessing and managing modifiable risk factors. Paediatr Drugs 2002;4(8):503-513. [Medline: 12126454]
15. Smith BA, Shuchman M. Problem of nonadherence in chronically ill adolescents: strategies for assessment and intervention. Curr Opin Pediatr 2005 Oct;17(5):613-618. [Medline: 16160536]
16. KyngAs HA, Kroll T, Duffy ME. Compliance in adolescents with chronic diseases: a review. J Adolesc Health 2000 Jun;26(6):379-388. [Medline: 10822178]
17. Hanghøj S, Boisen KA. Self-reported barriers to medication adherence among chronically ill adolescents: a systematic review. J Adolesc Health 2014 Feb;54(2):121-138. [doi: 10.1016/j.jadohealth.2013.08.009] [Medline: 24182940]
18. Kahana S, Drotar D, Frazier T. Meta-analysis of psychological interventions to promote adherence to treatment in pediatric chronic health conditions. J Pediatr Psychol 2008 Jul;33(6):590-611 [FREE Full text] [doi: 10.1093/jpepsy/jsm128] [Medline: 18192300]

19. Mistry N, Keepanasseril A, Wilczynski NL, Nieuwlaat R, Ravall M, Haynes RB, Patient Adherence Review Team. Technology-mediated interventions for enhancing medication adherence. J Am Med Inform Assoc 2015 Apr;22(e1):e177-e193. [doi: 10.1093/jamia/ocu047] [Medline: 25726568]
20. Pai Ahna LH, McGrady M. Systematic review and meta-analysis of psychological interventions to promote treatment adherence in children, adolescents, and young adults with chronic illness. J Pediatr Psychol 2014 Sep;39(8):918-931. [doi: 10.1093/jpepsy/jsu038] [Medline: 24952359]
21. McGrady ME, Ryan JL, Gutiérrez-Colina AM, Fredericks EM, Towner EK, Pai ALH. The impact of effective paediatric adherence promotion interventions: systematic review and meta-analysis. Child Care Health Dev 2015 Nov;41(6):789-802. [doi: 10.1111/cch.12271] [Medline: 26255643]
22. McGrady ME, Hommel KA. Targeting Health Behaviors to Reduce Health Care Costs in Pediatric Psychology: Descriptive Review and Recommendations. J Pediatr Psychol 2016 Sep;41(8):835-848. [doi: 10.1093/jpepsy/jsv083] [Medline: 26359311]
23. Dean AJ, Walters J, Hall A. A systematic review of interventions to enhance medication adherence in children and adolescents with chronic illness. Arch Dis Child 2010 Sep;95(9):717-723. [doi: 10.1136/adc.2009.175125] [Medline: 20522463]
24. Graves MM, Roberts MC, Rapoff M, Boyer A. The efficacy of adherence interventions for chronically ill children: a meta-analytic review. J Pediatr Psychol 2010 May;35(4):368-382 [FREE Full text] [doi: 10.1093/jpepsy/jsp072] [Medline: 19710248]
25. Lenhart A. Teens, Social Media & Technology Overview 2015. Washington, DC; 2015. Pew Research Center URL: http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/ [accessed 2016-11-04] [WebCite Cache ID 6pIxGjb7Q]
26. Smith A. U.S. Smartphone Use in 2015. Washington, DC; 2015. Pew Research Center URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2016-11-05] [WebCite Cache ID 6pIxWHxOV]
27. Badawy SM, Thompson AA, Liem RI. Technology Access and Smartphone App Preferences for Medication Adherence in Adolescents and Young Adults With Sickle Cell Disease. Pediatr Blood Cancer 2016 May;63(5):848-852. [doi: 10.1002/pbc.25905] [Medline: 26844685]
28. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and Adolescents: A Systematic Review of Adolescents' Use of Mobile Phone and Tablet Apps That Support Personal Management of Their Chronic or Long-Term Physical Conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]
29. Moher D, Liberati A, Tetzlaff J, Altman DG. 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]
30. Badawy S, Sino M, Kaviany S. Text messaging and smartphone app interventions to improve adherence in adolescents and young adults: a systematic review and meta-analysis.: University of York PROSPERO URL: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015025907 [accessed 2017-03-28] [WebCite Cache ID 6pIxyDO1D]
31. World Health Organization. Department of Child and Adolescent Health and Development. Geneva, Switzerland; 2001. The Second Decade: Improving Adolescent Health and Development URL: http://apps.who.int/iris/bitstream/10665/64320/1/WHO_FRH_ADH_98.18_Rev.1.pdf [accessed 2016-10-25] [WebCite Cache ID 6pIy53H64]
32. Higgins JPT, Green S. The Cochrane Collaboration. 2011. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 URL: http://www.cochrane-handbook.org [accessed 2016-08-05] [WebCite Cache ID 6pIyMSTKD]
33. Lipsey M, Wilson D. Practical meta-analysis (Applied Social Research Methods). Thousand Oaks, CA: Sage; 2001.
34. Boker A, Feetham HJ, Armstrong A, Purcell P, Jacobe H. Do automated text messages increase adherence to acne therapy? Results of a randomized, controlled trial. J Am Acad Dermatol 2012 Dec;67(6):1136-1142. [doi: 10.1016/j.jaad.2012.02.031] [Medline: 22521201]
35. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
36. Creary SE, Gladwin MT, Byrne M, Hildesheim M, Krishnamurti L. A pilot study of electronic directly observed therapy to improve hydroxyurea adherence in pediatric patients with sickle-cell disease. Pediatr Blood Cancer 2014 Jun;61(6):1068-1073. [doi: 10.1002/pbc.24931] [Medline: 24436121]
37. Dowshen N, Kuhns LM, Johnson A, Holoyda BJ, Garofalo R. Improving adherence to antiretroviral therapy for youth living with HIV/AIDS: a pilot study using personalized, interactive, daily text message reminders. J Med Internet Res 2012;14(2):e51 [FREE Full text] [doi: 10.2196/jmir.2015] [Medline: 22481246]
38. Estepp JH, Winter B, Johnson M, Smeltzer MP, Howard SC, Hankins JS. Improved hydroxyurea effect with the use of text messaging in children with sickle cell anemia. Pediatr Blood Cancer 2014 Nov;61(11):2031-2036. [doi: 10.1002/pbc.25177] [Medline: 25132074]
39. Fabbrocini G, Izzo R, Donnarumma M, Marasca C, Monfrecola G. Acne smart club: an educational program for patients with acne. Dermatology 2014;229(2):136-140. [doi: 10.1159/000362809] [Medline: 25011657]
40. 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]

41. Garofalo R, Kuhns LM, Hotton A, Johnson A, Muldoon A, Rice D. A Randomized Controlled Trial of Personalized Text Message Reminders to Promote Medication Adherence Among HIV-Positive Adolescents and Young Adults. AIDS Behav 2016 May;20(5):1049-1059. [doi: 10.1007/s10461-015-1192-x] [Medline: 26362167]
42. Hammonds T, Rickert K, Goldstein C, Gathright E, Gilmore S, Derflinger B, et al. Adherence to antidepressant medications: a randomized controlled trial of medication reminding in college students. J Am Coll Health 2015;63(3):204-208. [doi: 10.1080/07448481.2014.975716] [Medline: 25338175]
43. Louch G, Dalkin S, Bodansky J, Conner M. An exploratory randomised controlled trial using short messaging service to facilitate insulin administration in young adults with type 1 diabetes. Psychol Health Med 2013;18(2):166-174. [doi: 10.1080/13548506.2012.689841] [Medline: 22646659]
44. McKenzie RB, Berquist WE, Foley MA, Park KT, Windsheimer JE, Litt IF. Text Messaging Improves Participation in Laboratory Testing in Adolescent Liver Transplant Patients. J Particip Med 2015;7 [FREE Full text] [Medline: 26213633]
45. Miloh T, Annunziato R, Arnon R, Warshaw J, Parkar S, Suchy FJ, et al. Improved adherence and outcomes for pediatric liver transplant recipients by using text messaging. Pediatrics 2009 Nov;124(5):e844-e850 [FREE Full text] [doi: 10.1542/peds.2009-0415] [Medline: 19822583]
46. Mulvaney SA, Anders S, Smith AK, Pittel EJ, Johnson KB. A pilot test of a tailored mobile and web-based diabetes messaging system for adolescents. J Telemed Telecare 2012 Mar;18(2):115-118 [FREE Full text] [doi: 10.1258/jtt.2011.111006] [Medline: 22383802]
47. Ostojic V, Cvoriscec B, Ostojic SB, Reznikoff D, Stipic-Markovic A, Tudjman Z. Improving asthma control through telemedicine: a study of short-message service. Telemed J E Health 2005 Feb;11(1):28-35. [doi: 10.1089/tmj.2005.11.28] [Medline: 15785218]
48. Ting TV, Kudalkar D, Nelson S, Cortina S, Pendl J, Budhani S, et al. Usefulness of cellular text messaging for improving adherence among adolescents and young adults with systemic lupus erythematosus. J Rheumatol 2012 Jan;39(1):174-179. [doi: 10.3899/jrheum.110771] [Medline: 22089460]
49. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
50. McGrady ME, Ryan JL, Brown GA, Cushing CC. Topical Review: Theoretical Frameworks in Pediatric Adherence-Promotion Interventions: Research Findings and Methodological Implications. J Pediatr Psychol 2015 Sep;40(8):721-726 [FREE Full text] [doi: 10.1093/jpepsy/jsv025] [Medline: 25817881]
51. Badawy SM, Kuhns LM. Texting and Mobile Phone App Interventions for Improving Adherence to Preventive Behavior in Adolescents: A Systematic Review. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e50 [FREE Full text] [doi: 10.2196/mhealth.6837] [Medline: 28428157]
52. Badawy SM, Kuhns LM. Economic Evaluation of Text-Messaging and Smartphone-Based Interventions to Improve Medication Adherence in Adolescents with Chronic Health Conditions: A Systematic Review. JMIR Mhealth Uhealth 2016 Oct 25;4(4):e121 [FREE Full text] [doi: 10.2196/mhealth.6425] [Medline: 27780795]
53. Centre for Reviews Dissemination. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care (Internet). York, UK; 2009. URL: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf [accessed 2016-03-02] [WebCite Cache ID 6pJ0smh62]
54. Dubben H, Beck-Bornholdt H. Systematic review of publication bias in studies on publication bias. BMJ 2005 Aug 20;331(7514):433-434 [FREE Full text] [doi: 10.1136/bmj.38478.497164.F7] [Medline: 15937056]
55. Bartolucci A, Hillegass W. Overview, Strengths, and Limitations of Systematic Reviews and Meta-Analyses. In: Chiapelli F, editor. Evidence-Based Practice: Toward Optimizing Clinical Outcomes. Berlin: Springer-Verlag Berlin Heidelberg; 2010:17-33.
56. Badawy SM, Thompson AA, Lai J, Penedo FJ, Rychlik K, Liem RI. Health-related quality of life and adherence to hydroxyurea in adolescents and young adults with sickle cell disease. Pediatr Blood Cancer 2017 Jun;64(6). [doi: 10.1002/pbc.26369] [Medline: 27896936]
57. Badawy SM, Thompson AA, Penedo FJ, Lai J, Rychlik K, Liem RI. Barriers to hydroxyurea adherence and health-related quality of life in adolescents and young adults with sickle cell disease. Eur J Haematol 2017 Mar 17. [doi: 10.1111/ejh.12878] [Medline: 28306171]
58. Eysenbach G, Consort-EHEALTHGroup. 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]
59. Augestad KM, Berntsen G, Lassen K, Bellika JG, Wootton R, Lindsetmo RO, Study Group of Research Quality in Medical Informatics and Decision Support (SQUID). Standards for reporting randomized controlled trials in medical informatics: a systematic review of CONSORT adherence in RCTs on clinical decision support. J Am Med Inform Assoc 2012;19(1):13-21 [FREE Full text] [doi: 10.1136/amiajnl-2011-000411] [Medline: 21803926]

Abbreviations

CHCs: chronic health conditions
GRADE: Grades of Recommendation, Assessment, Development, and Evaluation
HRQOL: health-related quality of life
m-DOT: Mobile Direct Observed Therapy
PEF: peak expiratory flow
PEFM: peak expiratory flow meter
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
SMD: standardized mean differences

Edited by G Eysenbach; submitted 31.03.17; peer-reviewed by W Evans, L Dayer; comments to author 24.04.17; revised version received 28.04.17; accepted 29.04.17; published 15.05.17.

Please cite as:
Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM
Text Messaging and Mobile Phone Apps as Interventions to Improve Adherence in Adolescents With Chronic Health Conditions: A Systematic Review
JMIR Mhealth Uhealth 2017;5(5):e66
URL: http://mhealth.jmir.org/2017/5/e66/
doi:10.2196/mhealth.7798
PMID:28506955

Original Paper

North American Public Opinion Survey on the Acceptability of Crowdsourcing Basic Life Support for Out-of-Hospital Cardiac Arrest With the PulsePoint Mobile Phone App

Katie N Dainty[1,2*], PhD; Haris Vaid[3*], MSc; Steven C Brooks[4*], MPH, MD

[1]Rescu, Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, ON, Canada

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

[3]School of Medicine, Queen's University, Kingston, ON, Canada

[4]Department of Emergency Medicine, Queen's University, Kingston, ON, Canada

[*]all authors contributed equally

Corresponding Author:
Katie N Dainty, PhD
Rescu
Li Ka Shing Knowledge Institute
St Michael's Hospital
30 Bond Street
Toronto, ON, M5B 1W8
Canada
Phone: 1 6474482485
Email: daintyk@smh.ca

Abstract

Background: The PulsePoint Respond app is a novel system that can be implemented in emergency dispatch centers to crowdsource basic life support (BLS) for patients with cardiac arrest and facilitate bystander cardiopulmonary resuscitation (CPR) and automated external defibrillator use while first responders are en route.

Objective: The aim of this study was to conduct a North American survey to evaluate the public perception of the above-mentioned strategy, including acceptability and willingness to respond to alerts.

Methods: We designed a Web-based survey administered by IPSOS Reid, an established external polling vendor. Sampling was designed to ensure broad representation using recent census statistics.

Results: A total of 2415 survey responses were analyzed (1106 from Canada and 1309 from the United States). It was found that 98.37% (1088/1106) of Canadians and 96% (1259/1309) of Americans had no objections to PulsePoint being implemented in their community; 84.27% (932/1106) of Canadians and 55.61% (728/1309) of Americans said they would download the app to become a potential responder to cardiac arrest, respectively. Among Canadians, those who said they were likely to download PulsePoint were also more likely to have ever had CPR training (OR 1.7, 95% CI 1.2-2.4; *P*=.002); however, this was not true of American respondents (OR 1.0, 95% CI 0.79-1.3; *P*=.88). When asked to imagine themselves as a cardiac arrest victim, 95.39% (1055/1106) of Canadians and 92.44% (1210/1309) of Americans had no objections to receiving crowdsourced help in a public setting; 88.79% (982/1106) of Canadians and 84.87% (1111/1309) of Americans also had no objections to receiving help in a private setting, respectively. The most common concern identified with respect to PulsePoint implementation was a responder's lack of ability, training, or access to proper equipment in a public setting.

Conclusions: The North American public finds the concept of crowdsourcing BLS for out-of-hospital cardiac arrest to be acceptable. It demonstrates willingness to respond to PulsePoint CPR notifications and to accept help from others alerted by the app if they themselves suffered a cardiac arrest.

(JMIR Mhealth Uhealth 2017;5(5):e63) doi:10.2196/mhealth.6926

KEYWORDS

sudden cardiac death; surveys and questionnaires; cardiopulmonary resuscitation; PulsePoint; North America

Introduction

Background

Each year, more than 400,000 people have out-of-hospital cardiac arrest (OHCA) in the United States and Canada. Early cardiopulmonary resuscitation (CPR) and defibrillation are key links in the chain of survival for OHCA. Studies in specific high-risk community settings, such as casinos [1], airports [2], and aboard aircraft [3], have reported increased survival with the implementation of CPR and automated external defibrillator (AED) training, along with an organized response to cardiac arrest emergencies. Evidence from the "public access defibrillation (PAD) trial," which randomized 993 community units, such as apartment complexes and community centers, demonstrated that an organized lay-responder PAD program, including CPR training, can double the chance of survival for individuals who suffer OHCA [4]. Many historical approaches to increasing bystander resuscitation are limited by the fact that bystanders are chosen by circumstance, and not design. Many OHCA victims have probably died without receiving the benefit of bystander resuscitation, while trained, capable, and willing rescuers were nearby, but just out of sight.

Crowdsourcing Basic Life Support for OHCA

Crowdsourcing has been defined as the process of "obtaining needed services, ideas, or content by soliciting contributions from a large group of people, and especially from the online community, rather than from traditional employees or suppliers" [5]. The PulsePoint Respond mobile device application (app) is an example of a technology solution applying the concept of crowdsourcing to the problem of OHCA. The PulsePoint Respond mobile device app uses crowdsourcing to address the problem of OHCA. In communities where PulsePoint has been implemented, citizens can download PulsePoint Respond onto their mobile device from the Apple App Store or Google Play free of charge. Minimal contact information is collected on each user to ensure privacy and confidentiality. A major component of PulsePoint implementation involves a coordinated communications and public outreach strategy, encouraging CPR-trained individuals to download the PulsePoint Respond app onto their mobile devices. Messaging, which specifically targets individuals with CPR training (eg, paramedics, emergency medical technicians, firefighters, other health care providers, and CPR course graduates), is supported by resources from the PulsePoint Foundation and is generally coordinated by a communications professional within the hosting public safety agency. By using the global positioning system, mapping functionality of mobile devices, along with cardiac arrest location data provided by local 911 emergency call centers, the PulsePoint system can send directed cardiac arrest notifications to PulsePoint Respond users in close proximity (default 400 m) to the event. The notifications include the exact location of suspected cardiac arrest emergencies and registered AEDs [6].

Study Objective

PulsePoint engages community volunteers outside of the traditional professional emergency response as users. Although the majority of PulsePoint volunteers are off-duty health care providers [7], anyone in the community can download the app and become a PulsePoint responder. When new users download the app, they are asked to declare that they are CPR trained and willing to respond in an emergency. Designed as a true crowdsourcing solution to remove barriers to registration and encourage large numbers of "good samaritans" to participate, this self-declaration is not vetted. This design is consistent with the general understanding that any attempt as resuscitation during cardiac arrest, whatever the quality, is better than no attempt at all. The traditional response to 911 calls for emergency medical conditions has been limited to professional emergency personnel with validated CPR training. PulsePoint represents a departure from the status quo and has raised concerns about privacy and public safety among public safety agencies approached to consider implementation.

These are potential barriers to implementation of crowdsourcing solutions for OHCA. It is not clear whether the public perception of crowdsourcing for emergency response to cardiac arrest, specifically around privacy and safety issues, is consistent with the privacy and liability concerns that may be harbored by decision makers in public safety agencies. This valuable information could guide future research and implementation efforts. Accordingly, our study objective was to evaluate the public perceptions of the PulsePoint mobile app in the North American setting. More specifically, the aim of this research was to determine the level of public acceptance of crowdsourcing basic life support (BLS) for cardiac arrest and identify specific concerns with this strategy.

Methods

A Web-based public opinion survey was conducted within Canada and the United States in collaboration with an established polling vendor, IPSOS Reid. This study was approved by the Queen's University Health Sciences Research Ethics Board.

Setting

This study used the IPSOS eNation Canadian and US Online Omnibus survey platforms [8]. Data were collected through stratified random sampling of the more than 800,000-member Web-based panel and the Ampario sample source. The IPSOS Web-based panel is recruited and maintained using double and triple opt-in screening processes to ensure maximum return from an engaged and representative audience. The panel is updated regularly and non-responders are removed.

Study Sample

The study sample was designed to be nationally representative of the adult Canadian and American populations. Data were weighted on gender, age, region, and income, based on census information, to ensure that the sample's composition reflected that of the reference population. The Canadian sampling process included an additional sampling of French-speaking respondents in Canada to provide a base for analysis within that group. The US sampling process included an additional sample of Spanish-speaking respondents to provide a base for analysis within that group. The margin of error associated with this technique on a sample size of 1000 adults is <±3.1% relative

to the result that would be attained after polling the entire population, 19 times out of 20 [9].

Study Questionnaire

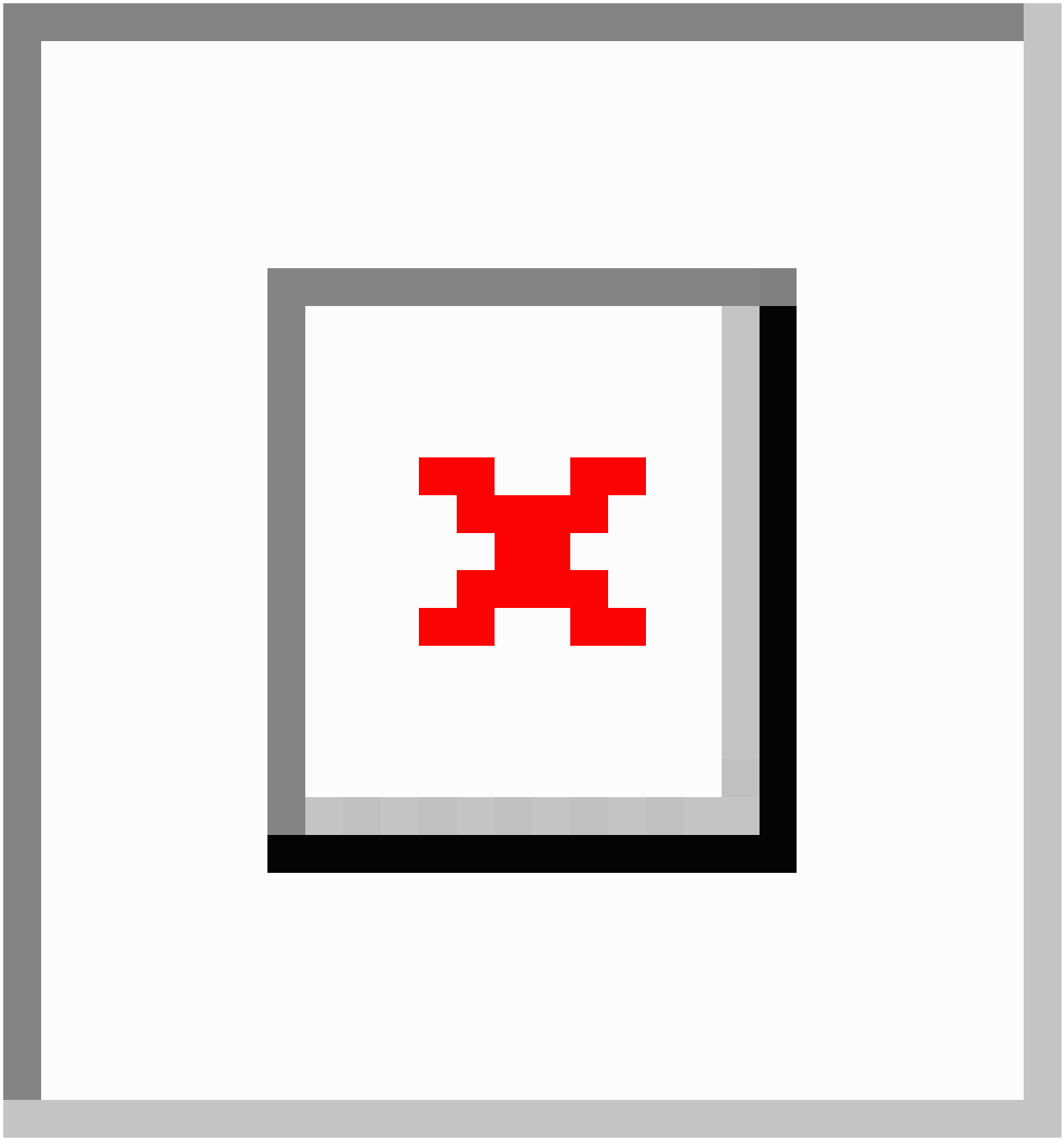
We presented the respondents with a short concept description of cardiac arrest and the PulsePoint app, followed by six closed-ended and four open-ended questions (Multimedia Appendix 1). Survey questions were developed by the research team and experts from IPSOS-Reid and pilot tested with a group of six lay public members affiliated with our research program. Our primary outcome was comfort level with the idea of having a PulsePoint-notified responder attend the respondent, imagining the scenario that they had suffered a sudden cardiac arrest. We specifically asked them about their comfort level when considering the situation of cardiac arrest in a public location versus a private (residential) location.

Data Analysis

Survey responses from the United States and Canada were treated independently because the samples were individually weighted according to US or Canadian census data. Analysis of the survey responses was carried out using descriptive statistics and comparison of groups based on responses to primary questions using chi-squared testing. We considered a *P* value of less than .05 to be statistically significant. We also conducted weighted-logistic regression analysis to assess the relationship between demographic factors (ie, the predictor variables) and two outcome variables (STATA version 13, STATA Corp). The weights used in our regression analysis were the same weights used to design the study sample, as described earlier. The outcome variables were the respondent's comfort level to crowdsourced BLS for an OHCA in a public location or a private location (see Questions 4 and 6 in Multimedia Appendix 1). Responses to the outcome variables, which were reported on a Likert scale, were dichotomized with specified cut-points set *a priori*. This was done to simplify the interpretation of the weighted-logistic regression analysis. Individuals who responded to the comfort-level questions with "very comfortable," "somewhat comfortable," or "neither comfortable nor uncomfortable" were coded as having "no objections." Individuals who responded with "somewhat uncomfortable" and "very uncomfortable" were coded as feeling "uncomfortable" (Figure 1).

A planned sensitivity analysis was conducted by using a different cut-point in the dichotomization of the outcome variables. Individuals who responded to the comfort-level questions with "very comfortable" and "somewhat comfortable" were coded as feeling "comfortable," while individuals who responded with "neither comfortable nor uncomfortable," "somewhat uncomfortable," and "very uncomfortable" were coded as feeling "less than comfortable" (Figure 1). The sensitivity analysis was conducted to test the stability of our primary regression models. We wanted to make sure that dichotomizing at a given cut-point, versus another cut-point, would not significantly influence the interpretation of our analysis. We built separate univariable and multivariable weighted logistic regression models for each outcome (comfort in private settings, comfort in public settings) in each national cohort (United States and Canada). Multivariable regression models included the following predictor variables: sex, education, household income, employment status, whether or not the respondent had children, age group, spoken language (Canadian data only), marital status (US data only), and race (US data only). We did not apply the finite population correction because our sample size was significantly smaller than the inference population. Open-ended, text-based responses regarding concerns were coded and summarized using standard content analysis [10].

Figure 1. Dichotomization of comfort level measures as reported by survey respondents for logistic regression analysis and sensitivity analysis.



Results

Data were collected between May 29 and June 3, 2015. A total sample of 2415 total surveys was collected: 1106 from Canada and 1309 from the United States. We did not calculate the response rate as the nature of the panel survey approach renders this statistic inapplicable. The demographic characteristics of the respondents are outlined in Table 1. A summary of the survey responses is presented in Table 2. At some point, 70% (769/1106) of Canadians and 63% (828/1309) of Americans had been trained in CPR. However, about one-third obtained their certification more than five years ago (United States: 441/1309, 34%; Canada: 356/1106, 32%). Among Canadians, those who said that they were likely to download PulsePoint were also more likely to have ever had CPR training (OR 1.7, 95% CI 1.2-2.4; *P*=.002). This suggested that knowledge of CPR might influence the likelihood of registering as a PulsePoint user. However, this relationship did not hold true among Americans (OR 1.0, 95% CI 0.79-1.3; *P*=.88).

Most Canadians (1088/1106, 98%) and Americans (1259/1309, 96%) had no objections to PulsePoint implementation in their community. Furthermore, most had no objections to receiving CPR if they were victim of cardiac arrest in a public location (Canada: 1055/1106, 95%; United States: 1210/1309, 92%) or private location (Canada: 982/1106, 89%; United States: 1111/1309, 85%).

Among Americans and Canadians, after adjusting for covariates, multivariable regression analyses found that demographic factors were not associated with comfort-level measures in the *public setting* scenario (Multimedia Appendices 2 and 3). Sensitivity analysis, with different cut-off points for the dichotomous outcome variables used in each model, revealed similar results. The multivariable regression analysis assessing comfort level to crowdsourced BLS in a *private setting* found that, compared with females, males had higher odds of having no objections to crowdsourced BLS (United States: OR 1.7, 95% CI 1.2-2.4; *P*=.004; Canada: OR 2.3, 95% CI 1.3-3.9) (Multimedia Appendices 2 and 3). Among Americans, compared with single individuals, married or co-habitating individuals had 1.7 times (95% CI 1.1-2.8; *P*.03) the odds of having no objections to crowdsourced BLS in a *private setting* scenario (Multimedia Appendix 2). Among Canadians, in comparison to individuals with household income <$50,000, those with income between $50,000 and $99,999 had half the odds of having no objections to crowdsourced BLS in a *private setting* scenario (OR 0.51, 95% CI 0.28-0.92, *P*=.02; Multimedia Appendix 3). However, this finding was not robust to the outcome cut-off selection; the odds ratio for this covariate was not statistically significant in the sensitivity analysis.

When asked, less than 60% (1421/2415) of the respondents identified any concerns with the PulsePoint mobile device app as described. The results are summarized in Figure 2. The most common concerns raised by the respondents were lack of training among PulsePoint responders and trust issues (Canada: 174/1106, 16%; United States: 241/1309, 18%). This was largely driven by concerns that PulsePoint responders might not have sufficient ability or training to provide effective assistance (Canada: 140/1106, 13%; United States: 190/1309, 15%). Some expressed security concerns, including legal liabilities, especially among Americans, and being stuck in a risky or dangerous situation. When asked to consider cardiac arrest in a public setting, concerns around lack of training and ability were the most significant here as well (Canada: 235/1106, 21%; United States: 350/1309, 27%). Security (Canada: 111/1106, 10%; United States: 215/1309, 16%) and privacy (Canada: 28/1106, 2%; United States: 9/1309, 1%) issues were less significant concerns. When asked to consider cardiac arrest in a private (ie, residential) setting, lack of training and ability remained the most significant concern (Canada: 165/1106, 15%; United States: 244/1309, 19%). Security (Canada: 131/1106, 12%; United States: 173/1309, 13%) and privacy issues (Canada: 129/1106, 12%; United States: 159/1309, 12%) were also generally more significant concerns in this setting.

Among Canadians and Americans, 90% (995/1106) and 89% (1158/1309) felt that it was important that responders have up-to-date CPR certification, respectively. This is even more pronounced among those who say they would download the app (Canada: OR 5.2, 95% CI 3.0-9.1; *P*<.001; United States: OR 4.3, 95% CI 2.8-6.6; *P*<.001).

Table 1. Respondent demographics.

Demographics	Canada (N=1106), n (%)	United States (N=1309), n (%)	*P* value
Gender			
Male	540 (49)	579 (44)	.05
Female	566 (51)	730 (56)	
Age (in years)			
18-34	315 (28)	381 (29)	.94
35-54	416 (38)	484 (37)	
55+	375 (34)	443 (34)	
Primary language			
English	838 (76)	1153 (88)	-
Spanish[a]	-[c]	156 (12)	
French	268 (24)	-	
Education			
Less than high school	71 (6)	49 (4)	<.001
High school diploma	210 (19)	277 (21)	
Post-secondary	466 (42)	475 (36)	
University degree	359 (32)	509 (39)	
Household income			
< $50,000	304 (27)	438 (33)	<.001
$50,000-99,000	335 (30)	602 (46)	
$100,000-149,000	252 (23)	179 (14)	
$150,000 +	72 (7)	90 (7)	
Marital status[b]			
Single	-	314 (24)	-
Married or cohabitating	-	799 (61)	
Widowed	-	61 (5)	
Divorced or separated	-	135 (10)	
Children			
Yes	315 (28)	347 (26)	.34
No	791 (72)	962 (74)	
Completed CPR training			
Ever	769 (70)	828 (63)	.003
Within the last year	128 (12)	154 (12)	<.001
In the past 5 years	286 (26)	233 (18)	
> 5 years ago	356 (32)	441 (34)	

[a]Data on the proportion of Spanish-speaking Canadians were not available for Canadian respondents.

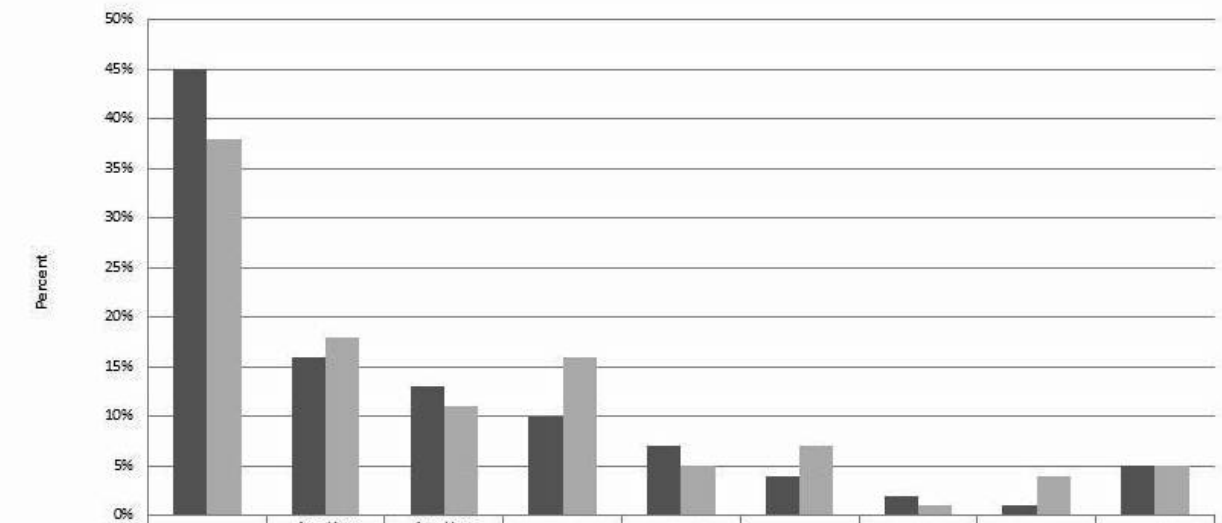
[b]Data on marital status were not available for Canadian respondents.

[c]Hyphens indicate that the data were not collected.

Table 2. Survey responses by question.

Question	Canada, n (%)	United States, n (%)	*P* value
To what extent do you agree or disagree that the PulsePoint Respond app is something that you would want to be made available in your community?			
	N=1106	N=1309	
Agree or strongly agree	903 (82)	1005 (77)	.002
Neither agree nor disagree	186 (17)	254 (19)	
Disagree or strongly disagree	18 (1)	50 (4)	
If the PulsePoint app was available in your community, how likely are you to download the app onto your mobile device?			
	N=932	N=1189	
Likely or very likely	583 (62)	728 (61)	.26
Neither likely nor unlikely	183 (20)	214 (18)	
Unlikely or very unlikely	166 (18)	247 (21)	
If you suffered a cardiac arrest in a *public setting* (eg, walking down the street, in a park, at the mall, at work), how comfortable would you be with nearby PulsePoint users being notified of your exact location and coming to help you until professional crews arrived?			
	N=1106	N=1309	
Comfortable or very comfortable	882 (80)	1002 (77)	.02
Neither comfortable nor uncomfortable	174 (16)	207 (16)	
Uncomfortable or very uncomfortable	51 (5)	99 (8)	
If you suffered a cardiac arrest in a *private setting* (eg, your home or the home of a friend or relative), how comfortable would you be with nearby PulsePoint users being notified of your exact location and coming to help you until professional crews arrived?			
	N=1106	N=1309	
Comfortable or very comfortable	796 (72)	892 (68)	.04
Neither comfortable nor uncomfortable	187 (17)	219 (17)	
Uncomfortable or very uncomfortable	124 (11)	198 (15)	
How important is it to you that all PulsePoint users, who could potentially be notified of cardiac arrest locations, have a valid and up-to-date CPR certification?			
	N=1106	N=1309	
Important or very important	995 (90)	1158 (89)	.54
Neither important nor unimportant	93 (8)	124 (9)	
Unimportant or very unimportant	18 (2)	27 (2)	

Figure 2. Concerns about the implementation of the Pulsepoint™ application by category.



	None	App User Ability/Training/ History	App User Availability/Timing	Safety/Security Concerns	Technologic Concerns	Victim Concerns	Privacy Concerns	Misuse Concerns	Other/Miscellaneous
CANADA (n=1106)	45%	16%	13%	10%	7%	4%	2%	1%	5%
USA (n=1309)	38%	18%	11%	16%	5%	7%	1%	4%	5%

Figure Legend:
App user ability/training/history: training, skills, confidence, stress managing difficult situation, trust, defer to trained health professionals, etc.
App user availability/timing: too inconvenient, awkward, reaching patient in time, guilt for not helping, proximity, finding person, too many helpers, etc.
Safety, security: legal liabilities, risky or dangerous situations
Technology concerns: effectiveness, accessibility, distrust in app, don't have smartphone
Victim concerns: injuring, hurting or losing the patient, stranger, patient family refuse help
Privacy concern: tracking and privacy
Misuse: use by criminals, scammers, fakers, and overuse
Others concerns: notifications, lack of awareness, and accuracy of information on app.

Discussion

Principal Findings

In a representative sample of the North American public, we found that most are in favor of crowdsourcing BLS for cardiac arrest in their community with the PulsePoint system. In addition, we found that most have no objections with the concept of receiving help from anonymous PulsePoint users in the setting of a public location cardiac arrest emergency, occurring outside of the hospital. When prompted, a minority of respondents raised concerns regarding issues around training and capability of PulsePoint responders, safety, and privacy. These results provide important insight for future implementation of such systems.

Crowdsourcing is increasingly being used to address challenges in health care research and delivery. Professional researchers crowdsource cohorts from health social networks for conducting traditional studies regularly so they can quickly and efficiently get opinions and test potential interventions [11-14]. For example, a new Web-based program, CrowdMed, aims to leverage the "wisdom of the crowd" by giving patients an opportunity to submit their cases of undiagnosed illnesses and interact with case solvers to obtain diagnostic possibilities [14]. Scientists from the University of Southampton and The University of Pennsylvania Perelman School of Medicine ran a "My Heart Map Challenge" to create a map of automatic external defibrillators (AEDs) in the city of Pittsburgh. The map was populated by members of the general public identifying and submitting photos and information about public AEDs [15].

Traditionally, 911 calls result in a response from vetted on-duty professional responders. PulsePoint and other similar systems represent a paradigm shift in the way in which we deploy help for 911 callers reporting a possible cardiac arrest and provide the ability to connect responders and victims in a timelier way to address the issue of low bystander response rates. Recently, Ringh et al found that rates of bystander-initiated CPR could be significantly increased with the use of a similar mobile phone positioning system. That system also located mobile phone users and dispatched lay volunteers who were trained in CPR to a patient nearby with OHCA [16]. Zijlstra et al and Pijls et al had similar findings in their studies of SMS text message (short message service, SMS) alert systems in Sweden and the Netherlands, respectively [17,18]. However, there are challenges with such systems, including low response rates, misidentification of cardiac arrest victims, and technical difficulties, such as excessive activation radii, and insufficient user density in the community [18]. As such, these systems have not received much scientific attention. During recent attempts to implement the PulsePoint system in several Canadian jurisdictions because of concerns over protecting privacy and

uncertainty about public opinion or potential backlash [7]. Our work demonstrates that public opinion strongly supports the implementation of PulsePoint-type apps and that many members of the public are willing to become responders, if given the opportunity.

Technologies are not separate from the society in which they are embedded, but are rather integral to the advancement of the social environment [19]. Society's increasing dependence on technologies, however, comes with an increased need to closely examine "society-technology" interactions. While on the one hand, a new technology may bring about radical changes in society, on the other hand, the fate of that technology rests with the society in which it is being applied. Much research has been conducted on risk and benefit perceptions and public attitudes, as these are believed to be the major factors influencing public acceptance of technologies [20-24]. In a review by Gupta et al on the sociopsychological determinants of public acceptance of technologies, perceived risk, perceived benefit, trust and culpability, knowledge, individual differences, and attitude are traditionally the most often reported or cited determinants [19]. This maps closely with the domains we included in our study. The most common concern identified with respect to PulsePoint implementation was a potential crowdsourced responders' lack of ability, training, knowledge, or having proper equipment in a public setting. In the case of PulsePoint, our results would indicate that the general public feels that the perceived benefit outweighs any perceived risk. Further, while issues of trust and privacy exist, in general, people would find it more than acceptable to use and to receive help from other users.

Findings from this work have several implications for the design of new crowdsourcing apps for cardiac arrest response or improvement of existing apps of this nature. For example, to address the concerns raised by respondents regarding the knowledge of responders, designers may consider including a training requirement to download the application. This could be as informal as agreeing to watch a short 5-minute video before download or as formal as uploading proof of professional CPR training. Each may have an impact on the public's willingness to download and would need to be monitored. Apps may also need to consider collecting further registration information from responders in order to ensure the safety of victims as well as to gather information on bystander response, something that is currently very difficult to collect.

Limitations

As with any survey-based study, our research has limitations. Our survey was conducted using a Web-based survey platform, which may create a bias toward individuals with computer literacy and access to the Internet. Due to the nature of the panel survey approach, we were unable to calculate the traditional response rate. Opinions may differ between responders and nonresponders to the survey, so our results may have been subject to selection bias. Finally, our weighted-logistic regression analysis was limited by a small sample size (within variables). This may limit the precision and power of our measures of association. Thus, although an association may not have been found in these analyses, a relationship between certain demographic characteristic and comfort level to crowdsourced BLS in a public or private setting may still exist due to type-two error.

Conclusions

A key conclusion of the American Heart Association statement on the use of mobile devices, social media, and crowdsourcing as digital strategies to improve emergency cardiovascular care is that there is a clear need for rigorous research to build the scientific evidence base for their effectiveness and safety [25]. Our findings provide the first empirical scientific evidence that the North American public supports the implementation of the PulsePoint mobile device application to crowdsource BLS for OHCA. The large majority of people are comfortable with the concept of receiving BLS from nearby PulsePoint users when given the hypothetical situations of suffering a cardiac arrest in a public or private location. Concerns around PulsePoint responders' abilities, level of training, and personal safety should be considered when planning implementation in communities as well as when designing future versions of the PulsePoint system. The results of our analysis should be very useful for decision-makers considering implementation of this strategy. Our findings can also guide developers of crowdsourcing solutions in addressing concerns related to trust and responder training in order to optimize community uptake.

Acknowledgments

This research was funded through a grant from the Laerdal Foundation for Acute Medicine.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey questionnaire.

[PDF File (Adobe PDF File), 172KB - mhealth_v5i5e63_app1.pdf]

Multimedia Appendix 2

Weighted regression analysis for the USA population.

[PDF File (Adobe PDF File), 34KB - mhealth_v5i5e63_app2.pdf]

Multimedia Appendix 3

Weighted regression analysis for the Canadian population.

[PDF File (Adobe PDF File), 36KB - mhealth_v5i5e63_app3.pdf]

References

1. Valenzuela TD, Roe DJ, Nichol G, Clark LL, Spaite DW, Hardman RG. Outcomes of rapid defibrillation by security officers after cardiac arrest in casinos. N Engl J Med 2000 Oct 26;343(17):1206-1209. [doi: 10.1056/NEJM200010263431701] [Medline: 11071670]
2. Caffrey SL, Willoughby PJ, Pepe PE, Becker LB. Public use of automated external defibrillators. N Engl J Med 2002 Oct 17;347(16):1242-1247. [doi: 10.1056/NEJMoa020932] [Medline: 12393821]
3. Page RL, Joglar JA, Kowal RC, Zagrodzky JD, Nelson LL, Ramaswamy K, et al. Use of automated external defibrillators by a U.S. airline. N Engl J Med 2000 Oct 26;343(17):1210-1216. [doi: 10.1056/NEJM200010263431702] [Medline: 11071671]
4. Hallstrom AP, Ornato JP, Weisfeldt M, Travers A, Christenson J, McBurnie MA, Public Access Defibrillation Trial Investigators. Public-access defibrillation and survival after out-of-hospital cardiac arrest. N Engl J Med 2004 Aug 12;351(7):637-646. [doi: 10.1056/NEJMoa040566] [Medline: 15306665]
5. Merriam-Webster. Definition of crowdsourcing URL: https://www.merriam-webster.com/dictionary/crowdsourcing [accessed 2017-04-10] [WebCite Cache ID 6pcjMusEl]
6. Pulsepoint. URL: http://www.pulsepoint.org [accessed 2017-01-11] [WebCite Cache ID 6nRMLgsxT]
7. Brooks SC, Simmons G, Worthington H, Bobrow BJ, Morrison LJ. The PulsePoint Respond mobile device application to crowdsource basic life support for patients with out-of-hospital cardiac arrest: Challenges for optimal implementation. Resuscitation 2016 Jan;98:20-26. [doi: 10.1016/j.resuscitation.2015.09.392] [Medline: 26475397]
8. IPOSOS. IPSOS Canadian Omnibus Services URL: http://www.ipsos.ca/en/products-tools/public-affairs/omnibus-surveys-product/canadian-omnibus-services.aspx [accessed 2017-01-11] [WebCite Cache ID 6nRMEhDGt]
9. ipsosglobaladvisor. IPSOS Global Advisor URL: http://www.ipsosglobaladvisor.com/methodology.aspx [accessed 2017-01-11] [WebCite Cache ID 6nRM7AfTF]
10. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
11. Gottlieb A, Hoehndorf R, Dumontier M, Altman RB. Ranking adverse drug reactions with crowdsourcing. J Med Internet Res 2015;17(3):e80 [FREE Full text] [doi: 10.2196/jmir.3962] [Medline: 25800813]
12. Kristan J, Suffoletto B. Using online crowdsourcing to understand young adult attitudes toward expert-authored messages aimed at reducing hazardous alcohol consumption and to collect peer-authored messages. Transl Behav Med 2015 Mar;5(1):45-52 [FREE Full text] [doi: 10.1007/s13142-014-0298-4] [Medline: 25729452]
13. Chang T, Verma BA, Shull T, Moniz MH, Kohatsu L, Plegue MA, et al. Crowdsourcing and the accuracy of online information regarding weight gain in pregnancy: a descriptive study. J Med Internet Res 2016 Apr 07;18(4):e81 [FREE Full text] [doi: 10.2196/jmir.5138] [Medline: 27056465]
14. Meyer AN, Longhurst CA, Singh H. Crowdsourcing diagnosis for patients with undiagnosed illnesses: an evaluation of CrowdMed. J Med Internet Res 2016;18(1):e12 [FREE Full text] [doi: 10.2196/jmir.4887] [Medline: 26769236]
15. Merchant RM, Asch DA, Hershey JC, Griffis HM, Hill S, Saynisch O, et al. A crowdsourcing innovation challenge to locate and map automated external defibrillators. Circ Cardiovasc Qual Outcomes 2013 Mar 01;6(2):229-236 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.113.000140] [Medline: 23481522]
16. Ringh M, Rosenqvist M, Hollenberg J, Jonsson M, Fredman D, Nordberg P, et al. Mobile-phone dispatch of laypersons for CPR in out-of-hospital cardiac arrest. N Engl J Med 2015 Jun 11;372(24):2316-2325. [doi: 10.1056/NEJMoa1406038] [Medline: 26061836]
17. Zijlstra J, Beesems S, De Haan RJ, Koster R. Psychological impact on dispatched local lay rescuers performing bystander cardiopulmonary resuscitation. Resuscitation 2015 Jul;92:115-121. [doi: 10.1016/j.resuscitation.2015.04.028] [Medline: 25957944]
18. Pijls RW, Nelemans P, Rahel B, Gorgels AP. A text message alert system for trained volunteers improves out-of-hospital cardiac arrest survival. Resuscitation 2016 Aug;105:182-187 [FREE Full text] [doi: 10.1016/j.resuscitation.2016.06.006] [Medline: 27327230]
19. Mendelson J, Berg M. Getting the right bystander to the right place at the right time. Resuscitation 2016 Oct;107:A11-A12. [doi: 10.1016/j.resuscitation.2016.07.002] [Medline: 27496258]
20. Fischer ARH, De Vries PW. Everyday behaviour and everyday risk: An exploration how people respond to frequently encountered risks. Health Risk Soc 2008;10:397.
21. Otway HJ, Von Winterfeldt D. Beyond acceptable risk: on the social acceptability of technologies. Policy Sci 1982;14(3):247-256. [doi: 10.1007/BF00136399]

22. Siegrist M, Keller C, Kastenholz H, Frey S, Wiek A. Laypeople's and experts' perception of nanotechnology hazards. Risk Anal 2007 Feb;27(1):59-69. [doi: 10.1111/j.1539-6924.2006.00859.x] [Medline: 17362400]
23. Sjöberg L. Attitudes toward technology and risk: going beyond what is immediately given. Policy Sci 2002;35(4):379-400. [doi: 10.1023/A:1021354900928]
24. Sjöberg L, Fromm J. Information technology risks as seen by the public. Risk Anal 2001 Jun;21(3):427-441. [Medline: 11572424]
25. Gupta N, Fischer ARH, Frewer LJ. Socio-psychological determinants of public acceptance of technologies: a review. Public Underst Sci 2012;21(7):782-795. [doi: 10.1177/0963662510392485]
26. Rumsfeld JS, Brooks SC, Aufderheide TP, Leary M, Bradley SM, Nkonde-Price C, American Heart Association Emergency Cardiovascular Care Committee; Council on Cardiopulmonary, Critical Care, Perioperative and Resuscitation; Council on Quality of Care and Outcomes Research; Council on Cardiovascular and Stroke Nursing; Council on Epidemiology and Prevention. Use of mobile devices, social media, and crowdsourcing as digital strategies to improve emergency cardiovascular care: a scientific statement from the American Heart Association. Circulation 2016 Aug 23;134(8):e87-e108. [doi: 10.1161/CIR.0000000000000428] [Medline: 27334603]

Abbreviations

AED: automated external defibrillator
AEDs: automatic external defibrillators
BLS: basic life support
CPR: cardiopulmonary resuscitation
OHCA: out-of-hospital cardiac arrest
PAD: public access defibrillation

Edited by G Eysenbach; submitted 14.11.16; peer-reviewed by M Leary, O Anya; comments to author 28.12.16; revised version received 27.01.17; accepted 22.02.17; published 17.05.17.

Please cite as:
Dainty KN, Vaid H, Brooks SC
North American Public Opinion Survey on the Acceptability of Crowdsourcing Basic Life Support for Out-of-Hospital Cardiac Arrest With the PulsePoint Mobile Phone App
JMIR Mhealth Uhealth 2017;5(5):e63
URL: http://mhealth.jmir.org/2017/5/e63/
doi:10.2196/mhealth.6926
PMID:28526668

Original Paper

Uptake of an Incentive-Based mHealth App: Process Evaluation of the Carrot Rewards App

Marc Mitchell[1,2*], PhD; Lauren White[2*], MSc; Paul Oh[1], MD; David Alter[1], MD, PhD; Tricia Leahey[3], PhD; Matthew Kwan[4], PhD; Guy Faulkner[5], PhD

[1]University Health Network, Toronto, ON, Canada
[2]Carrot Insights Inc, Toronto, ON, Canada
[3]University of Connecticut, Storrs, CT, United States
[4]McMaster University, Hamilton, ON, Canada
[5]University of British Columbia, Vancouver, BC, Canada
[*]these authors contributed equally

Corresponding Author:
Marc Mitchell, PhD
University Health Network
347 Rumsey Road
Toronto, ON,
Canada
Phone: 1 416 660 7881
Fax: 1 416 425 0301
Email: mmitc26@uwo.ca

Abstract

Background: Behavioral economics has stimulated renewed interest in financial health incentives worldwide. The Carrot Rewards app was developed as part of a public-private partnership to reward Canadians with loyalty points (eg, movies and groceries) for downloading the app, referring friends, and completing an average of 1 to 2 educational health quizzes per week ("micro-learning"), with long-term objectives of increasing health knowledge and encouraging healthy behaviors.

Objective: The main objective of this study was to evaluate uptake of a loyalty points-based mHealth app during the exclusive 3-month launch period in British Columbia (BC), Canada. The secondary aims were to describe the health and sociodemographic characteristics of users, as well as participation levels (eg, proportion of quizzes completed and friends referred).

Methods: The app was promoted via loyalty program email campaigns (1.64 million emails). Number of downloads and registrations (users enter age, gender, and valid BC postal code to register) were collected. Additional sociodemographics were inferred by linking postal codes with census data at the local health area (LHA) level. Health risk assessments were also deployed. Participation levels were collected over 3 months and descriptive data were presented.

Results: In 3 months, 67,464 individuals downloaded the app; in its first week, Carrot Rewards was the most downloaded health app in Canada. Among valid users (n=57,885; at least one quiz completed), the majority were female (62.96%; 36,446/57,885) and aged 18 to 34 years (54.34%; 31,459/57,885). More than half of the users (52.40%; 30,332/57,885) resided in LHAs where the median personal income was below the provincial average (Can $28,765). Furthermore, 64.42% (37,291/57,885) of users lived in metropolitan (ie, urban) LHAs, compared with 56.17% of the general BC population. The most prevalent risk factors were "not" meeting physical activity guidelines (72.70%; 31,765/43,692) and "not" getting the flu shot last year (67.69%; 30,286/44,739). Regarding participation, 60.05% (34,761/57,885) of users were classified as "very high" engagers (>75% quiz completion rate).

Conclusions: Early results suggest that loyalty points may promote mHealth app uptake. The app was downloaded by younger females especially, and BC residents from higher and lower income regions were equally represented. Loyalty points appear to have driven participation throughout the inaugural 3-month period (ie, quiz completion).

(JMIR Mhealth Uhealth 2017;5(5):e70) doi:10.2196/mhealth.7323

KEYWORDS

financial incentives; mHealth; behavioral economics; public health

Introduction

The cost of health care in Canada is rising at an unsustainable rate [1]. Although health care costs are shared between federal, provincial, and territorial governments in Canada, ultimately provinces and territories are responsible for health care spending, organization, delivery, and management of health care services and incur a greater proportion of health care costs. About 40% of provincial and territorial budgets are spent on health care alone, and total health expenditure growth in 2015 was forecasted to be 1.6% [2]—this, at a time when the number of older Canadians at high-risk of developing chronic conditions is on the rise (5 million Canadians aged 65 years+ in 2011 vs 10 million expected in 2036) [3]. Expensive chronic diseases, however, are not just reserved for older adults, as approximately half of Canadians aged 20 years and above live with at least one chronic condition [4]. In the 2014 economic analysis, the costliest modifiable risk factors in Canada were smoking, physical inactivity, and obesity, accounting for Can $50 billion in direct health care expenses [5]. Although improving an individual's risk factor profile is possible, health behavior change can be extremely difficult. Persisting levels of smoking (18%) [6], physical inactivity (80%) [7], and overweight or obesity (54%) [8] in Canada provides case-in-point. Acknowledging that health behaviors are influenced across multiple domains and require multipronged and multisectoral solutions, governments are looking for new, innovative, and collaborative ways of promoting healthy lifestyles [9,10].

Behavioral economics, a branch of economics complimented by insights from psychology, has stimulated renewed interest in using financial incentives to motivate healthy behaviors on a population scale [11-13]. Briefly, behavioral economics recognizes that human decisions are biased in systematic ways and that these "decision biases" can make it hard for people to make self-beneficial choices [14,15]. For example, people often succumb to the "present bias" when making health-related decisions because they place disproportionate emphasis on the present "costs" of health behaviors (ie, time out of a busy schedule) and discount the future benefits (eg, quality of life). According to behavioral economics, increasing the immediately rewarding aspects of health behaviors (with a financial incentive, for example) may offset the so-called "present bias," increasing the likelihood of action [14,15]. A growing evidence base seems to support this theoretical rationale. For instance, several systematic reviews and meta-analyses have shown that incentives generally stimulate health behaviors [16-22]. Despite some notable gaps in the literature (eg, long-term effects are not clear), financial health incentives have grown in popularity. The best examples of broad-based application come from Germany and the United States where incentive-based public health policies have been in place since 2004 and 2014, respectively [11,23]. In the United States, the Patient Protection and Affordable Care Act (2010) allows employers (as of the year 2014) to reimburse their employees up to US $1500 per year for engaging in healthy behaviors or reaching health targets [23]. An often-cited limitation of these programs and policies, however, is the notorious delay between (1) behavior or outcome achievement and (2) reward [15,23,24]. Generally speaking, this lag has been too long (and the incentives not large or meaningful enough) to elicit the desired behavioral responses or health outcomes [23,24].

The pervasive use of smartphones in Canada (in 2015, 73% of Canadian adults owned a smartphone—a relative increase of 7% compared with 2014) [25], presents governments with an opportunity to offer financial health incentives on a population scale, with little delay between behavior and reward—leveraging people's predictable tendency to overvalue the benefits they experience in the present. Moreover, smartphone capabilities have evolved (eg, accelerometry, global positioning system [GPS]), now allowing for the tracking of a range of objectively measured health behaviors (eg, walking and flu clinic visits). This may increase intervention effectiveness as rewards tied to objectively measured behaviors tend to work better than rewards contingent on self-reported ones [16]. Finally, there are several program design features that can be manipulated to optimize incentive effectiveness (eg, timing, type, magnitude, probability, and schedule), and loyalty points (ie, points given by retailers to promote customer loyalty) have emerged as a promising new incentive "type" (loyalty points vs cash, vouchers or health insurance premium reimbursements or discounts) [26-28]. Not only are Canadians avid loyalty point collectors (90% members of one or more loyalty programs) [29], but the perceived value of their loyalty points may be inflated (in part, because it is not clear how much a single point is actually worth), lowering the reward magnitude needed to stimulate behaviors [16,30-32]. These intervention components may be particularly appealing for governments looking to deploy incentives as efficiently as possible.

The Carrot Rewards app is a new multisectoral mHealth initiative that harnesses the pervasiveness of smartphones and Canadians' affinity to loyalty programs to reward healthy behaviors. Developed as part of a public-private partnership [33], one of the distinguishing features of the app is its ability to immediately reward users with loyalty points. Specifically, the app rewards Canadians with loyalty points (to go to the movies or grocery store, for gas or travel) for downloading the app, referring friends, and completing an average of 1 to 2 short educational health quizzes each week ("micro-learning"), with the ultimate goal of increasing health knowledge and promoting healthy behaviors. The primary objective of this process evaluation, the first step of a multi-stage evaluation, is to determine the uptake (eg, downloads, completed registrations) of the new app during its 3-month launch period in British Columbia, Canada. The secondary study objectives are to describe the sociodemographic and health behavior characteristics of Carrot Rewards' users as well as participation levels (eg, proportion of quizzes completed, number of friends referred). A qualitative analysis of app store reviews is also presented.

Methods

App Overview

Background

Carrot Insights Inc is a private company that developed the Carrot Rewards app in partnership with the Public Health Agency of Canada. British Columbia was the company's founding provincial partner (the federal-provincial funding arrangement is described elsewhere) [33]. Carrot Insights Inc partnered with 4 major Canadian loyalty programs to offer a variety of popular incentives (ie, points can be redeemed for groceries, air travel, movies, or gas). The company also partnered with 4 Canadian health charities to assist with the development and approval of educational health quiz content (ie, Heart & Stroke Foundation of Canada, Diabetes Canada, YMCA Canada, and the British Columbia Healthy Living Alliance). The marketing assets of 4 loyalty and 1 charity partners were also leveraged such that in the initial weeks of the app launching in British Columbia, partners sent 1.64 million emails to their members (of which 800,167 could be "tracked," representing the email campaigns of 3 out of the 5 partners).

Registration

Carrot Rewards was made available on the iTunes and Google Play app stores on March 3, 2016 in both English and French (Canada's official languages). Upon downloading the app, users were asked to enter their age, gender, postal code, and loyalty program card number to complete registration (the card of their choice: for either the movies, gas, groceries, or travel loyalty program). To successfully register, users must have entered a valid BC postal code and be 13 years or older (age cut-off of participating loyalty programs). After registration was completed (and a reward worth about US $0.74 earned), a unique promotional code was provided to each user. This code could be shared with "friends;" if a "friend" downloaded the app using the unique code, both parties received bonus points (again worth US $0.74). British Columbians could download the app in one of three ways: organically (ie, finding it in the app store on their own), via partner email invitation, or by using the promotional code "friend referral" mechanism.

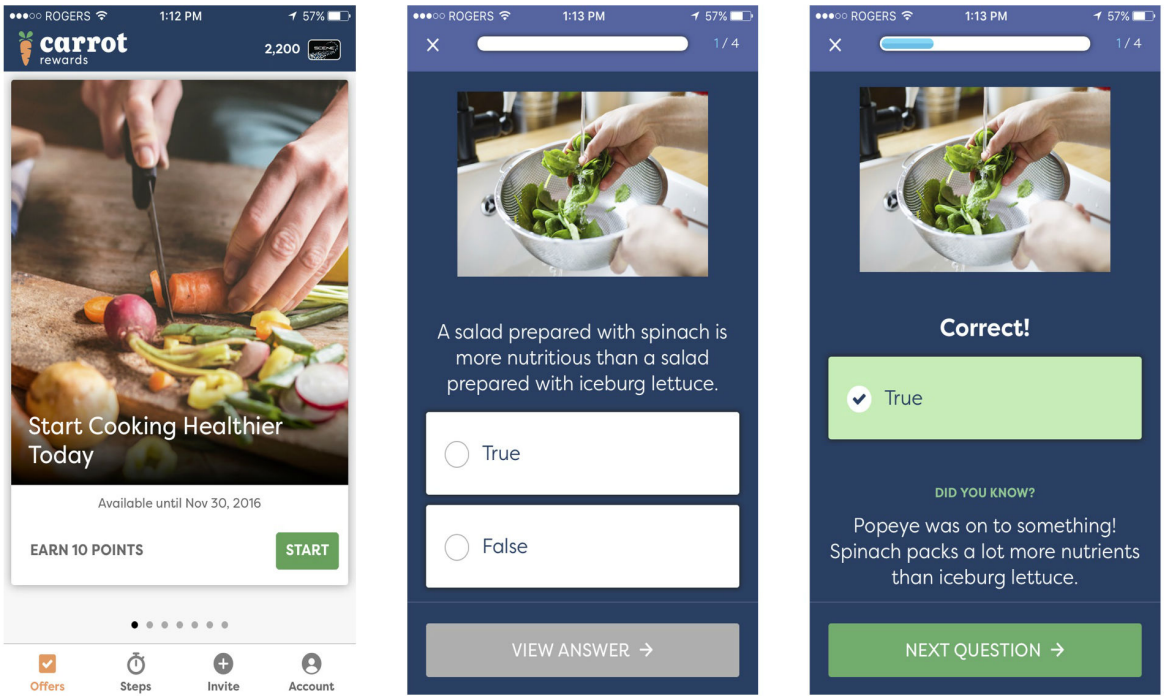
Intervention

Once the app was downloaded and registration was completed, users were "offered" an average of 1 to 2 educational health quizzes per week, each containing 5 to 7 questions related to healthy eating, physical activity and sedentary behavior, smoking, low-risk drinking, mental health, and immunization—public health priorities identified by the BC Ministry of Health (see colorful and visually appealing screenshots in Figure 1 and quiz "schedule" in Multimedia Appendix 1). In addition, quizzes were developed to inform and familiarize users about self-regulatory health skills or "stepping stone" behaviors (ie, goal setting, tracking, action planning, and barrier identification), skills that have promoted health behaviors in the past [34]. After completing every question in a health quiz and immediately earning incentives (US $0.04 to US $1.48 depending on the length and timing of quiz, ie, earlier quizzes were worth more to stimulate program interest), users could view relevant health information on partner websites. Each health quiz was designed to take approximately 1-3 min to complete. Notably, quizzes were "released" in campaigns (each campaign included about 4 quizzes, so that only users who completed the first quiz in a campaign received subsequent quizzes). In addition to health quizzes, users could earn more points for completing separate health risk assessments (HRAs; Multimedia Appendix 2, which included items from national health surveys (regarding physical activity, eating and smoking habits, alcohol consumption, mental health and overall well-being, as well as frequency of influenza immunization). HRAs were made available in the first 4 weeks of the program (ie, "Carrot Health Survey, 1" and "Carrot Health Survey, 2"). For a full description of the Carrot Rewards intervention, including incentive design features, see Multimedia Appendices 3 and 4. Notable intervention changes during the 3-month evaluation period include gradual reductions in (1) quiz frequency (from about 15 per month to 5 per month, depending on the date a user downloaded the app) as well as (2) reward magnitude (reduction to about 5% of original reward value). Quiz frequency and reward magnitudes were set high initially in an attempt to maximize interest and early participation; both were reduced over time to ensure intervention spending did not outpace the finite budget.

Theoretical Rationale

Whereas the Carrot Rewards app is grounded in behavioral economics, broader theoretical considerations regarding how rewards motivate human behaviors may further improve the effectiveness of this approach. Self-determination theory (SDT) is a global theory of motivation that focuses on the extent to which behaviors are controlled by external agents (eg, physicians) or contingencies (eg, rewards) [34]. Whereas behavioral economics describes how rewards can be used to exploit the "present bias" and be a catalyst for change, SDT describes the conditions under which rewards may promote quality, sustained change [34-36]. It was hypothesized that an SDT-informed approach to knowledge building (eg, focus on enjoyment and self-regulatory skills such as goal setting) may help users "internalize" the reasons to engage in healthy behaviors, and in doing so, increase the potential for longer-term change (see Multimedia Appendix 5 for an overview of SDT and its application).

Figure 1. Carrot Rewards app health quiz screenshots (quiz title, question and answer page examples; 'link out' screen not shown).



Procedure

A flowchart is presented in Figure 2 outlining the flow of users through the program: from download to registration to valid user (ie, completed the initial quiz). "Incomplete" users did not successfully complete registration. "Excluded" users are those without valid age and gender data, or who had managed to create more than one account. "Registered" users successfully completed the registration. "Inactive" users did not complete the initial health quiz and therefore did not receive subsequent quizzes. Finally, "valid" users successfully completed the initial health quiz and received subsequent quizzes. App usage data was also collected during the inaugural 3-month period in British Columbia. Research ethics board (REB) approval was not required as the University Health Network REB did not consider this project as research, as described in the Tri-Council Policy Statement V.2, and therefore, did not fall under their purview. The University Health Network REB retrospectively issued an ethics waiver letter (#16-0129) on December 22, 2016. Additionally, as part of the Carrot Rewards' privacy policy (agreed to by users upon registration), users were informed that data collected in the app for reporting purposes would only be done at the aggregate or deidentified level.

Data Collection

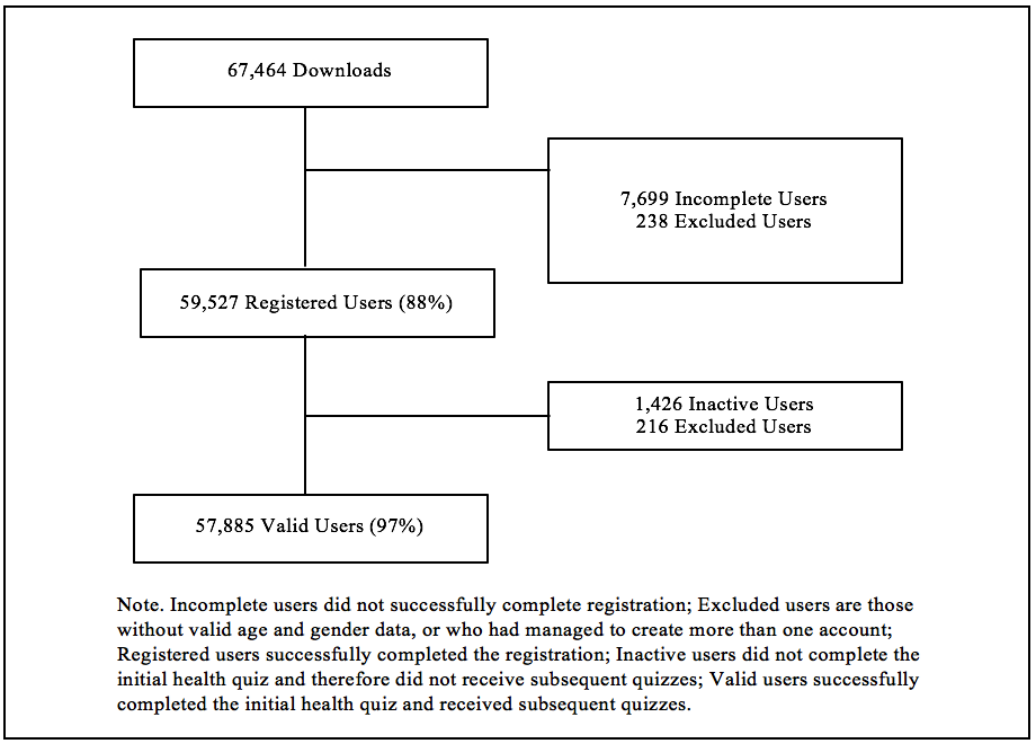
Sociodemographic and Health Behavior Characteristics

In addition to providing information on age and gender, other sociodemographic information was inferred by linking user postal codes with census data (ie, National Household Survey and British Columbia Geographic Service Areas data) at the local health area (LHA) level—there are 89 LHAs in British Columbia. Specifically, median personal income, postsecondary educational attainment, proportion of population identifying as a visible minority, and population density data (metropolitan, urban-rural, rural, and remote) based on LHA were matched to individual users. Users could also complete the HRAs for points. HRA items were modified from valid and reliable questionnaires (eg, Canadian Community Health Survey; Canadian Health Measures Survey) to fit the 15-word app limit (see Multimedia Appendix 2 for HRA items and their source).

Participation

App usage data included: individual-level quiz completion data, number of quiz "link outs" clicked, and number of successful friend referrals. Users' participation levels were classified in the following way: low (<25% of quizzes completed), medium (25- 49% completed), high (50- 74% completed), or very high (>75% completed). Aggregate "app uninstall" data is also reported.

Figure 2. Flowchart of Carrot Rewards app users.



Statistical Analyses

Descriptive statistics for sociodemographic (self-reported and inferred from postal code) and health behavior variables are provided. Users with missing age and gender data were excluded from this analysis (n=454). SPSS version 19.0 was used for all statistical analyses.

Results

Uptake

Within the first 3 months, 67,464 individuals downloaded the app (62.10% iTunes, Apple Inc Cupertino, CA; 37.90% Google Play, Google Inc Mountain View, CA). In its first week, the Carrot Rewards app was the most downloaded health app in Canada, although only being made available to the population of British Columbia, which makes up about 13% of the national population [37]. During the launch week, 13.79% of the "trackable" emails promoting the app were opened (110,383/800,167) and 1.20% initiated a "click through" to one of the two app stores (19,339/1,610,167; "click throughs" could be tracked for most emails). The majority of downloads occurred organically (46.00%; 31,033/67,464) or via friend referral (31.80%; 21,454/67,464). Among those who downloaded the app, 59,527 (88.23%) individuals successfully completed the registration (ie, registered users). Among registered users, 57,885 (97.24%) were classified as "valid," having completed at least the initial heath quiz. The Carrot Rewards uninstall rate (data available for Android devices only) was 17.00% (4,518/26,576) during the 3-month study period.

Sociodemographic and Health Characteristics

Among valid users (n=57,885), the majority were female (62.96%) and aged 18 to 34 years (54.34%; Table 1). More than half of users (52.40%) resided in LHAs where the median personal income was below the provincial average (Can $28,765). Sixty-four percent of users lived in a metropolitan area (LHAs with at least 190,001 residents), compared with 56.17% of the general BC population. The most prevalent chronic disease risk factors were "not" meeting physical activity guidelines (72.70% vs 78% of Canadians in general; Multimedia Appendix 2) and not getting the influenza vaccine in the past year (67.69% vs 69% of Canadians in general; Multimedia Appendix 2). Nine percent of users reported smoking, compared with 14% of all British Columbians (Multimedia Appendix 2).

Table 1. Sociodemographics of Carrot Rewards users and British Columbians in general.

Sociodemographics	Carrot Rewards n (%)	British Columbia n (%)
Gender	N=57,885	N=4,683,139[a]
Female	36,446 (62.96)	2,358,862 (50.37)
Male	20,179 (34.86)	2,324,277 (49.63)
Prefer not to answer	1,044 (1.80)	Not applicable
Other	216 (0.37)	Not applicable
Age (years)	N=57,885	N=4,683,139[a]
13-17	1,391 (2.40)	249,064 (6.08)
18-24	11,954 (20.65)	443,889 (10.84)
25-34	19,505 (33.69)	634,709 (15.50)
35-44	11,642 (20.11)	611,064 (14.92)
45-54	7,625 (13.17)	683,346 (16.69)
55-64	4,182 (7.22)	653,766 (15.96)
65-79	1,502 (2.59)	605,065 (14.77)
80-99	84 (0.15)	212,999 (5.20)
Income	N=57,885	N=3,722,755[b]
Users in LHAs[f] where the personal median income is below the BC average, n (%)	30,332 (52.40)	1,939,170 (52.09)
Postsecondary education	N=57,885	N=2,029,760[c]
Users in LHAs with fewer residents than average with a postsecondary education	28,788 (49.73)	926,550 (45.64)
Visible minority	N=57,885	N=1,180,845[d]
Users in LHAs with fewer residents than average identifying as a visible minority, n (%)	49,854 (86.13)	951,220 (80.55)
Population density	N=57,885	N=4,624,660[e]
Users living in LHAs categorized by population density, n (%)		
Remote	296 (0.51)	67,257 (1.45)
Rural	2272 (3.93)	496,865 (10.74)
Urban or rural	18,026 (31.14)	1,462,866 (31.63)
Metropolitan	37,291 (64.42)	2,597,672 (56.17)

[a]Values based on British Columbia's 2015 population estimates.

[b]Values based on British Columbia's 2011 National Household Survey. BC personal median income is Can $28,765.

[c]Values based on British Columbia's 2011 National Household Survey. The proportion of the BC population who reported a postsecondary education is 56%.

[d]Values based on British Columbia's 2011 National Household Survey. The proportion of the BC population who reported as a visible minority is 38%.

[e]Values based on 2012, 2016 British Columbia Geographic Service Areas data. Remote is classified as a population of 0-10,000 people; rural as 10,001-40,000 people; urban or rural as 40,001-190,000 people, and metro as 190,001+.

[f]LHAs: local health areas.

When asked "What is (are) your biggest health priority(s) in the next 6 months?" in a week 8 quiz, the most commonly cited priorities were: increase physical activity (71.00%; 26,239/36,956), improve eating habits (67.97%; 25,119/36,956), and manage stress levels (41.50%; 15,340/36,956). Finally, when users were asked whether they had learned anything in the first 8 weeks of the program, 93.99% (35,008/37,243) replied "yes."

Participation

In the first 3 months, 15 health quizzes and both HRAs were sent to users. A total of 879,616 quizzes or HRAs were sent and 690,111 were completed (78.45%; Multimedia Appendix 1). Regarding quiz completion rate, 60.17% (34,834) of valid users were classified as very high engagers (>75% quiz completion rate), 13.75% (7,954) as high (51-75%), 7.01% (4,057) as medium (26-50%), and 18.45% (10,675) as low (<25%) (N=57,885). Forty-three percent of users (24,870/57,885)

completed all 15 quizzes. Regarding attrition, an examination of quiz completion rates for the first quizzes of campaigns 2 (64.53% at week 5; 37,243/57,708) and 3 (62.53% at week 10; 35,855/57,338) suggest participation levels persisted, at least in the short-term. Upon quiz completion, users could learn more about each topic by viewing additional Web-based health resources. Users "clicked out" for more information about 4% of the time (26,574 "clicks;" Multimedia Appendix 6). On average, users sent 3.5 email referrals. The acceptance rate for email referrals was 20.35% (17,540/86,194). Twenty-seven percent (9694/36,445) and 6.08% (2216/36,445) of users successfully referred at least one friend, via email and promotional code, respectively.

Qualitative Analysis

A conventional content analysis [38] of iTunes and Google Play app store reviews was conducted to examine written reviews. All reviews posted from March 3 to June 6 were compiled and categorized by one researcher as either a positive or negative comment, or both (some reviews had both "positive" and "negative" comments). Both iTunes (n=66; 38.8%) and Google Play (n=104; 61.2%) store reviews were examined. Among "positive" comments (n=119), users highly enjoyed receiving loyalty points for participating in the program (31.9%), users liked learning new health information (11.7%), and users described the app as simple and easy to use (6.7%). The majority of "negative" comments (n=141) were regarding problems loading the app (eg, frozen screens, issues with entering contact information; 15.6%), the reduction in loyalty points over time (11.34%; necessary due to budgetary constraints), and referring friends to the app (9.9%; users required to remember and manually enter friend email addresses). At the 3-month mark (end of the study period), the average app store rating was 2.9 stars for the iTunes store and 3.8 stars for the Google Play store, based on 5-point scale rating systems.

Discussion

Principal Findings

Despite much promise, the public health potential of financial health incentives has not yet been realized [20,24]. In particular, incentive programs have been limited in their ability to scale and accommodate entire populations (not just for employees with extended health insurance, for example). Pervasive smartphone use and efficient loyalty points-based incentives allow for broader implementation that is not prohibitively costly. This study represents the first in a planned series with a focus on immediate objectives. Intermediate (eg, health knowledge improvements, short-term improvement in physical activity) and longer term (eg, sustained increase in physical activity, social return of investment) objectives will be evaluated at a later date. The main finding in this study was that an mHealth app that rewards users with loyalty points for downloading (and engaging with) the app was readily downloaded. Uptake was high despite only being available in one province. It is believed that the combination of a comprehensive email campaign, the promise of loyalty points that BC residents already use, and the idea of being rewarded to get or stay healthy, drove interest and uptake. Because there was no control group, the isolated effects of the incentives cannot be established. The early results from this program seem to be enhanced by the use of loyalty points, but not without recognizing the effect of other intervention characteristics (such as the private-public partnership, massive marketing effort, visually appealing design, and so on).

The Carrot Rewards app was the most downloaded health app in Canada during its launch week, despite it being available only in one province. Whereas 46% of downloads occurred "organically," the 2-way referral bonus (if user successfully refers friend and friend downloads app, both get bonus points) increased the number of users, representing 32% of all downloads. Importantly, 88% of users who downloaded the app successfully registered (eg, entered their loyalty card number), suggesting the onboarding procedure was not onerous. The quiz completion rate (78%) was also higher than expected, with 43% of users completing all 15 of the available "quizzes." In contrast, a similar mHealth app that used loyalty points to increase downloads but not participation found that 85% of its users were categorized as "very low" or "low" engagers in the first month (ie, fewer than 15 completed "challenges"—though the daily "ask" from the user in this case was heavier with quizzes available every day for 30 days) [27]. Likewise, an eHealth platform promoted using loyalty points (for completing an HRA and enrolling) determined that less than 2% of the approximately 42,000 enrollees were using the tool 6 weeks later [26]. These results are not surprising since attrition is a hallmark of eHealth and mHealth interventions [31]; however, the results presented here suggest that modest incentives in the form of loyalty points may help drive engagement (at least over the course of 3 months). In addition, the number of "clicks" to Carrot Rewards partner websites (n=26,574) highlight the potential role of driving traffic to partner resources.

Whereas the app only required users to report age and gender to minimize friction during registration, supplementary sociodemographic information inferred from postal codes suggests that the current sample is broadly representative of British Columbians in general. Specifically, Carrot Rewards' captured users in both lower and higher income areas as well as in metropolitan and more rural areas across the province. These finding are consistent with the literature suggesting that lower income adults are especially sensitive to incentive interventions and likely to respond by signing-up for health interventions they may not have otherwise [39-41]. The Carrot Rewards' user base—in the first 3 months—was predominantly female and between the ages of 18 and 34 years. The incentive-based eHealth and mHealth interventions mentioned earlier report similar demographic profiles, with women especially being more likely to adopt and engage with these interventions (68% and 74% female, respectively) [26,27]. To attract more men and older adults to the platform, several approaches have been recommended, including conducting interviews and focus groups to learn more about intervention components that might appeal to these harder to reach groups (eg, recruitment techniques, loyalty program offerings, and behaviors targeted) [42]. Varying recruitment methods (essentially broadening the "entry points" beyond just email marketing and friend referral) to involve the health care system and specifically leaders within the system (eg, physicians) in

the recruitment process (for instance, as a chronic disease self-management tool), may be one way of broadening appeal to under-represented, higher-risk subpopulations.

There are several practical implications of this study. First, though some argue that target groups may disagree with the incentive approach, citing that governments, for example, should not be paying their citizens to engage in healthy behaviors (ie, opportunity cost, paternalism—"nanny state" concerns) [43,44], little evidence of this was uncovered in the quantitative or qualitative aspects of this study. Additionally, the app store ratings at the end of the evaluation period were lower than expected (initial and current ratings are higher—eg, 4+ stars), suggesting that the drop in health quiz frequency and reward magnitude, as well as early technical issues (eg, frozen screens), may have antagonized users. In the future, quiz frequencies and reward magnitudes should be set at levels that can be titrated up (if budgets allow, for example) rather than down (or intermittent reenforcement should be used). It is worth noting that Apple and Google Play store ratings differed (2.9 vs 3.8 out of 5, respectively), possibly because app refinements (eg, bug fixes) occurred at different rates for the iOS and Android versions of the app. Next, the app was successfully launched in large part because of the public-private partnerships that initiated its development. In particular, leveraging the marketing and health content assets of private and government or charity sector partners, respectively, while offering a new way for these partners to communicate with the general population, created a win-win situation for the parties involved. Lessons learned in the multiple-stakeholder development and delivery of the app will be applied as the app prepares to launch in other Canadian provinces or territories and in other sectors (eg, financial literacy). Whereas the preliminary results show that the use of financial incentives encourages engagement with a healthy behavior app supported via a public-private partnership, it does not provide evidence to support the notion that governments should necessarily sponsor such programs. Another practical implication of this work may be in demonstrating that smartphones can be used as vehicles for "immediate" health-related feedback and rewards on a population-scale and that loyalty points in particular may be a useful incentive "form," given that (1) users are already familiar with the loyalty programs (not a new "currency"), (2) users tend to overvalue these points (when compared with actual dollars) [16,30-32], and (3) users collect points in a variety of ways (not just by using the app), and so are happy for the app to contribute to their growing rewards "pool." Others looking to maximize the efficiency of their health incentive interventions may be able to apply some of the lessons learned here. The main theoretical contribution of this study may be that incentives need not be prohibitively large or costly to stimulate behaviors on a population-scale; this intervention aspect might actually protect against the often-cited risk of incentive intervention, which is that they undermine intrinsic motives, particularly when targeting health behaviors, rather than just "stepping stone" ones (eg, education) [34-36]. In this case, incentives were deployed not to provide a controlling function (eg, if incentives are very large) but rather an informative one (eg, feedback in the form of modest rewards and health information with every interaction), and may serve to support rather than "crowd out" motivation.

Limitations

This study is not without limitations. This analysis includes data from British Columbia only, and thus may not be generalizable to other regions that may join in the future. Also, data for this study was extracted before potentially fraudulent user accounts were deactivated. Regarding promotional code use (for referring friends), it was impossible to track number of times the unique codes were shared, as they could be shared on any number of social media sites (or other ways); only number of times promotional codes were "redeemed" were reported. App "uninstall" data was available for Android device users only, which represented only a 38% share of all devices used to download the app. Whereas the sociodemographic data inferred from users' postal codes is valuable to characterize the user base, strong conclusions cannot be drawn regarding these important sociodemographics. Furthermore, whereas HRA items were adapted from national surveys, their psychometric properties were not tested before implementation; there is evidence of concurrent validity though, with physical activity, flu shot, smoking, and mental health responses aligning with provincial and/or national statistics. The fruit and vegetable items in particular may have been subject to "anchor bias" where users tend to avoid extreme responses (eg, consuming 0 or 5 green veggies yesterday). In the current context, this may have led to overreporting of the number of times fruits and vegetables were consumed yesterday. The overall quiz completion rate (78%) may also have been inflated as users must have completed the first quiz in a "campaign" (a "campaign" is usually 4 quizzes long) in order to receive subsequent quizzes; so, completion rates for quizzes 2-4 in a campaign for example may be inflated since nonadherers do not receive them. Regarding "clicks" to partner websites, whereas 4% of the time users clicked for more information, it is not clear how much time users spent on these sites. Further examination of session times on partner sites is likely required. Whereas the qualitative analysis followed Hsieh and Shannon's (2005) published framework, it was only conducted by 1 researcher which may limit objectivity of the findings. At most, the result of this qualitative analysis is concept development. Further work is required to increase confidence that conclusions accurately portray the data (eg, employing qualitative data analysis software, peer debriefing). Whereas this study did not examine the impact of quiz frequency and reward magnitude reductions on participation levels over time, future research should explore this issue as well.

Future Directions

In the future, longitudinal analyses will be conducted to examine the impact of the intervention on changes in self-reported health behaviors as well as health knowledge (baseline vs follow-up). Studies exploring healthy living resource awareness (eg, helpline, Web-resources), self-regulatory skill practices (eg, goal setting, tracking), and objectively measured behaviors (eg, personalized walking goals, getting the flu shot) are planned as well since these are more proximal to the behavioral, health and/or health care system outcomes of interest. A more comprehensive understanding of the contextual factors (eg,

demographics) and program features (eg, reward size, probability), as it relates to incentive program effectiveness, would be useful in informing program design in the future. In addition, qualitative work to ascertain intervention components that would appeal to harder to reach subgroups are needed, as are an exploration of new approaches to recruiting users. This study provides further evidence that incentives may be used to stimulate health-related behaviors, though more work is required to elucidate the conditions under which incentives can be used to drive "longer-term" changes. Longer-term cohort studies and other research designs are needed that can attribute behavioral, health, and health economic outcomes to incentive-based mHealth interventions within complex systems of health care.

Conclusions

The Carrot Rewards app has started to address the problem of scaling up incentive programs while maintaining fidelity to behavioral economics ("present bias;" little delay between behavior and reward for maximum effectiveness). Early results suggest that loyalty points may be used to promote the uptake of an mHealth app in a sample that broadly reflects the sociodemographic and health behavior profiles of British Columbians. A major challenge in mHealth is to develop innovative personalized interventions that can help individuals "maintain" healthy lifestyles and this should be the focus of future work. Their effects on behavioral outcomes (eg, steps per day, flu shot) should be explored, given recent advances in smartphone capabilities. Smartphones in general offer the potential to gather large amounts of data that can be used to better inform interventions. Moving forward, rich datasets should be used to drive the sustained changes needed to produce clinically and economically significant health outcomes.

Acknowledgments

The Carrot Rewards initiative has been made possible in part through funding from the Public Health Agency of Canada. The views expressed herein do not necessarily represent the views of the Public Health Agency of Canada or the British Columbia Ministry of Health. The authors also thank staff from the Public Health Agency of Canada, the British Columbia Ministry of Heath, the British Columbia Healthy Living Alliance, Heart & Stroke Foundation of Canada, Diabetes Canada, and YMCA Canada for providing expert health content advice over the course of this project. The Public Health Agency of Canada and the British Columbia Ministry of Health reviewed the manuscript.

Conflicts of Interest

Dr Marc Mitchell reports grant support from the Canadian Institutes of Health Research, the University Health Network, Green Shield Canada Inc, as well as in-kind research support from Cookson James Loyalty Inc. Furthermore, he reports consulting income from Carrot Insights Inc and stock options in Carrot Insights Inc. Lauren White is a Carrot Insights Inc employee and also reports stock options in Carrot Insights Inc. Dr Guy Faulkner is supported by a Canadian Institutes of Health Research-Public Health Agency of Canada (CIHR-PHAC) Chair in Applied Public Health. The other coauthors report no conflicts of interest.

Multimedia Appendix 1

Carrot Rewards app quiz schedule and completion rates.

[PDF File (Adobe PDF File), 173KB - mhealth_v5i5e70_app1.pdf]

Multimedia Appendix 2

Health behaviour characteristics of Carrot Rewards users.

[PDF File (Adobe PDF File), 192KB - mhealth_v5i5e70_app2.pdf]

Multimedia Appendix 3

Description of the Carrot Rewards app using the behavioral intervention technology (BIT) model.

[PDF File (Adobe PDF File), 165KB - mhealth_v5i5e70_app3.pdf]

Multimedia Appendix 4

Incentive Intervention Framework to describe the Carrot Rewards incentive program.

[PDF File (Adobe PDF File), 42KB - mhealth_v5i5e70_app4.pdf]

Multimedia Appendix 5

Application of self-determination theory in the development of the Carrot Rewards app.

[PDF File (Adobe PDF File), 151KB - mhealth_v5i5e70_app5.pdf]

Multimedia Appendix 6

Carrot Rewards app end-of-quiz partner website “link out” or “click.”.

[PDF File (Adobe PDF File), 169KB - mhealth_v5i5e70_app6.pdf]

References

1. Canadian Institute for Health Information. 2011. Health care cost drivers URL: https://secure.cihi.ca/free_products/health_care_cost_drivers_the_facts_en.pdf [accessed 2016-10-31] [WebCite Cache ID 6qPDmMFMo]
2. Canadian Institute for Health Information. 2015. National health expenditure trends URL: https://www.cihi.ca/en/spending-and-health-workforce/spending/national-health-expenditure-trends [accessed 2016-10-31] [WebCite Cache ID 6qPDpSuvr]
3. Statistics Canada. 2015. Canada's population estimates URL: http://www.statcan.gc.ca/daily-quotidien/150929/dq150929b-eng.htm [accessed 2016-10-31] [WebCite Cache ID 6qPDswPpZ]
4. Public Health Agency of Canada. 2016. Improving health outcomes: a paradigm shift URL: http://www.phac-aspc.gc.ca/cd-mc/assets/pdf/ccdp-strategic-plan-2016-2019-plan-strategique-cpmc-eng.pdf [accessed 2017-05-12] [WebCite Cache ID 6qPJpMRYT]
5. Krueger H, Turner D, Krueger J, Ready AE. The economic benefits of risk factor reduction in Canada: tobacco smoking, excess weight and physical inactivity. Can J Public Health 2014 Mar 18;105(1):e69-e78. [Medline: 24735700]
6. Statistics Canada. 2016 Mar 07. CANSIM lifestyle and social condition table URL: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/health73b-eng.htm [accessed 2016-10-31] [WebCite Cache ID 6qPEBaVU7]
7. Statistics Canada. 2016 Mar 7. CANSIM lifestyle and social condition table URL: http://www.statcan.gc.ca/pub/82-625-x/2015001/article/14135-eng.htm [accessed 2016-10-31] [WebCite Cache ID 6qPEJHyvm]
8. Statistics Canada. 2016 Mar 17. CANSIM lifestyle and social condition table URL: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/health81b-eng.htm [accessed 2016-10-31] [WebCite Cache ID 6qPEN4ovG]
9. Public Health Agency of Canada. 2013. Preventing chronic disease strategic plan 2013-2016 URL: http://publications.gc.ca/collections/collection_2014/aspc-phac/HP35-39-2013-eng.pdf [accessed 2017-05-12] [WebCite Cache ID 6qPEVHtZo]
10. Conference Board of Canada. 2015. Behavioural economics: nudging toward a culture of wellness URL: http://www.conferenceboard.ca/e-library/abstract.aspx?did=6739 [accessed 2017-05-12] [WebCite Cache ID 6qPEbTunZ]
11. Stock S, Schmidt H, Büscher G, Gerber A, Drabik A, Graf C, et al. Financial incentives in the German Statutory Health Insurance: new findings, new questions. Health Policy 2010 Jun;96(1):51-56. [doi: 10.1016/j.healthpol.2009.12.015] [Medline: 20106543]
12. Ries N. Financial incentives for weight loss and healthy behaviours. Healthc Pap 2012;7(3):5.
13. Schmidt H, Stock S, Doran T. Commonwealth Fund. 2012. Moving forward with Wellness Incentives Under the Affordable Care Act: lessons from Germany URL: http://www.commonwealthfund.org/publications/issue-briefs/2012/may/moving-forward-with-wellness-incentives-under-the-aca [accessed 2017-05-12] [WebCite Cache ID 6qPFBvw72]
14. Camerer CF, Loewenstein G. Advances in behavioural economics. In: Behavioural Economics: Past, Present, Future. New Jersey: Princeton University Press; 2003.
15. Loewenstein G, Asch DA, Volpp KG. Behavioral economics holds potential to deliver better results for patients, insurers, and employers. Health Aff (Millwood) 2013 Jul;32(7):1244-1250 [FREE Full text] [doi: 10.1377/hlthaff.2012.1163] [Medline: 23836740]
16. Mitchell MS, Goodman JM, Alter DA, John LK, Oh PI, Pakosh MT, et al. Financial incentives for exercise adherence in adults: systematic review and meta-analysis. Am J Prev Med 2013 Nov;45(5):658-667. [doi: 10.1016/j.amepre.2013.06.017] [Medline: 24139781]
17. Strohacker K, Galarraga O, Williams DM. The impact of incentives on exercise behavior: a systematic review of randomized controlled trials. Ann Behav Med 2014 Aug;48(1):92-99 [FREE Full text] [doi: 10.1007/s12160-013-9577-4] [Medline: 24307474]
18. Burns R, Donovan AS, Ackermann RT, Finch EA, Rothman AJ, Jeffery RW. A theoretically grounded systematic review of material incentives for weight loss: implications for interventions. Ann Behav Med 2012 Dec;44(3):375-388. [doi: 10.1007/s12160-012-9403-4] [Medline: 22907712]
19. Haff NM, Patel R, Lim R, Zhu J, Troxel AB, Asch DA, et al. The role of behavioral economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviors: a meta-analysis. Am J Health Promot 2015;29(5):314-323. [doi: 10.4278/ajhp.140714-LIT-333] [Medline: 25928816]
20. Mantzari E, Vogt F, Shemilt I, Wei Y, Higgins J, Marteau TM. Personal financial incentives for changing habitual health-related behaviors: a systematic review and meta-analysis. Prev Med 2015 Jun;75:75-85 [FREE Full text] [doi: 10.1016/j.ypmed.2015.03.001] [Medline: 25843244]
21. Cahill K, Perera R. Competitions and incentives for smoking cessation. Cochrane Database Syst Rev 2011 Apr 13(4):CD004307. [doi: 10.1002/14651858.CD004307.pub4] [Medline: 21491388]

22. Finkelstein EA, Haaland BA, Bilger M, Sahasranaman A, Sloan RA, Nang EE, et al. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. Lancet Diabetes Endocrinol 2016 Dec;4(12):983-995. [doi: 10.1016/S2213-8587(16)30284-4] [Medline: 27717766]
23. Volpp KG, Asch DA, Galvin R, Loewenstein G. Redesigning employee health incentives--lessons from behavioral economics. N Engl J Med 2011 Aug 04;365(5):388-390 [FREE Full text] [doi: 10.1056/NEJMp1105966] [Medline: 21812669]
24. Patel MS, Asch DA, Troxel AB, Fletcher M, Osman-Koss R, Brady J, et al. Premium-based financial incentives did not promote workplace weight loss in a 2013-15 study. Health Aff (Millwood) 2016 Jan;35(1):71-79. [doi: 10.1377/hlthaff.2015.0945] [Medline: 26733703]
25. Canadian Radio-television and Telecommunications Commission (CRTC). Communications Monitoring Report 2016 URL: http://www.crtc.gc.ca/eng/publications/reports/PolicyMonitoring/2016/cmr.pdf [accessed 2017-05-12] [WebCite Cache ID 6qPFuRh3E]
26. Liu S, Hodgson C, Zbib A, Payne A, Nolan RP. The effectiveness of loyalty rewards to promote the use of an internet-based heart health program. J Med Internet Res 2014 Jul 2;16(7):e163 [FREE Full text] [doi: 10.2196/jmir.3458]
27. Goyal S, Morita PP, Picton P, Seto E, Zbib A, Cafazzo JA. Uptake of a consumer-focused mHealth application for the assessment and prevention of heart disease: the <30 days study. JMIR Mhealth Uhealth 2016 Mar 24;4(1):e32 [FREE Full text] [doi: 10.2196/mhealth.4730] [Medline: 27012937]
28. Richards M, Sindelar J. Rewarding healthy food choices in SNAP: behavioral economic applications. Milbank Quarterly 2013 Jun 13;91(2):395-412. [doi: 10.1111/milq.12017]
29. Yahoo Inc. Talking Loyalty. 2015 Mar 24. URL: http://yahooadvertisingca.tumblr.com/post/114494275160/new-study-explores-how-loyalty-programs-drive
30. Mitchell M, Goodman JM, Alter DA, Oh PI, Faulkner GE. 'Will walk for groceries': acceptability of financial health incentives among Canadian cardiac rehabilitation patients. Psychol Health 2014;29(9):1032-1043. [doi: 10.1080/08870446.2014.904863] [Medline: 24731024]
31. Mitchell MS, Faulkner GE. On supplementing “Foot in the door” incentives for eHealth program engagement. J Med Internet Res 2014 Jul 25;16(7):e179 [FREE Full text] [doi: 10.2196/jmir.3701] [Medline: 25092221]
32. Maritz Insights. 2012 Feb 7. The loyalty report URL: http://www.maritzcastage.com/studio/MLM/LoyaltyReport/MaritzInsightsTheLoyaltyReport2012.pdf [accessed 2017-05-12] [WebCite Cache ID 6qPIJ9Ycd]
33. Government of Canada. 2015 Jul 29. National Healthy Living Platform URL: http://news.gc.ca/web/article-en.do?nid=1010069 [accessed 2017-05-12] [WebCite Cache ID 6qPIP09Kk]
34. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol Bull 1999 Nov;125(6):627-668. [Medline: 10589297]
35. Mitchell M, Faulkner G. A “nudge” at all? The jury is still out on financial health incentives. Healthc Pap 2012;12(4):31. [Medline: 23713398]
36. 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;9:78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
37. Statistics Canada. 2016 Sep 28. CANSIM population estimate and projection table URL: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo02a-eng.htm [accessed 2017-05-12] [WebCite Cache ID 6qPIcNfDJ]
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. John L, Loewenstein G, Volpp K. Empirical observations on longer- term use of incentives for weight loss. Prev Med 2012;55(Suppl):S68-S74. [doi: 10.1016/j.ypmed.2012.01.022.?]
40. Haff N, Patel M, Lim R. The role of behavioural economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviours: a meta-analysis. Am J Hlth Promo 2015 May;29(5):314-323.
41. Finkelstein E, Brown D, Brown D, Buchner D. A randomized study of financial incentives to increase physical activity among sedentary older adults. Prev Med 2008 Aug;47(2):182-187. [doi: 10.1016/j.ypmed.2008.05.002] [Medline: 18571226]
42. Giles EL, Sniehotta FF, McColl E, Adams J. Acceptability of financial incentives and penalties for encouraging uptake of healthy behaviours: focus groups. BMC Public Health 2015 Jan 31;15:58-64 [FREE Full text] [doi: 10.1186/s12889-015-1409-y] [Medline: 25636330]
43. Park J, Metlay J, Asch JM, Asch DA. The New York Times readers' opinions about paying people to take their medicine. Health Educ Beh 2012 Mar 29;39(6):725-731. [doi: 10.1177/1090198111428645]
44. Park J, Mitra N, Asch DA. Public opinion about financial incentives for smoking cessation. Prev Med 2012 Nov;55 Suppl:S41-S45. [doi: 10.1016/j.ypmed.2012.06.013] [Medline: 22735040]

Abbreviations

HRA: health risk assessment
LHA: local health area
REB: research ethics board

SDT: self-determination theory

Edited by G Eysenbach; submitted 19.01.17; peer-reviewed by J Cafazzo, IO Adepoju, S Goyal; comments to author 16.02.17; revised version received 17.03.17; accepted 03.04.17; published 30.05.17.

Please cite as:
Mitchell M, White L, Oh P, Alter D, Leahey T, Kwan M, Faulkner G
Uptake of an Incentive-Based mHealth App: Process Evaluation of the Carrot Rewards App
JMIR Mhealth Uhealth 2017;5(5):e70
URL: http://mhealth.jmir.org/2017/5/e70/
doi:10.2196/mhealth.7323
PMID:28559224

Original Paper

Ownership and Use of Commercial Physical Activity Trackers Among Finnish Adolescents: Cross-Sectional Study

Kwok Ng[1], MSc, PhD; Jorma Tynjälä[1], PhD; Sami Kokko[1], PhD

Research Centre for Health Promotion, Faculty of Sport and Health Sciences, University of Jyvaskyla, Jyväskylä, Finland

Corresponding Author:
Kwok Ng, MSc, PhD
Research Centre for Health Promotion
Faculty of Sport and Health Sciences
University of Jyvaskyla
PO Box 35 (L)
Jyväskylä, 40014
Finland
Phone: 358 451499919
Fax: 358 451499919
Email: kwok.ng@jyu.fi

Abstract

Background: Mobile phone apps for monitoring and promoting physical activity (PA) are extremely popular among adults. Devices, such as heart rate monitors or sports watches (HRMs/SWs) that work with these apps are at sufficiently low costs to be available through the commercial markets. Studies have reported an increase in PA levels among adults with devices; however, it is unknown whether the phenomena are similar during early adolescence. At a time when adolescents start to develop their own sense of independence and build friendship, the ease of smartphone availability in developed countries needs to be investigated in important health promoting behaviors such as PA.

Objective: The objective of this study was to investigate the ownership and usage of PA trackers (apps and HRM/SW) among adolescents in a national representative sample and to examine the association between use of devices and PA levels.

Methods: The Finnish school-aged physical activity (SPA) study consisted of 4575 adolescents, aged 11-, 13-, and 15-years, who took part in a web-based questionnaire during school time about PA behaviors between April and May 2016. Binary logistic regression analyses were used to test the associations between moderate to vigorous physical activity (MVPA) and devices, after controlling for gender, age, disability, and family affluence.

Results: PA tracking devices have been categorized into two types, which are accessible to adolescents: (1) apps and (2) HRM/SW. Half the adolescents (2351/4467; 52.63%) own apps for monitoring PA, yet 16.12% (720/4467) report using apps. Fewer adolescents (782/4413; 17.72%) own HRM/SW and 9.25% (408/4413) use HRM/SW. In this study, users of HRM/SW were 2.09 times (95% CI 1.64-2.67), whereas users of apps were 1.4 times (95% CI 1.15-1.74) more likely to meet PA recommendations of daily MVPA for at least 60 min compared with adolescents without HRM/SW or without apps.

Conclusions: To our knowledge, this is the first study that describes the situation in Finland with adolescents using PA trackers and its association with PA levels. Implications of the use of apps and HRM/SW by adolescents are discussed.

(JMIR Mhealth Uhealth 2017;5(5):e61) doi:10.2196/mhealth.6940

KEYWORDS

social determinants of health; mobile phone; health promotion; disabled children; physical activity; adolescent

Introduction

According to national statistics in Finland, by the age of 16 years, all Finnish adolescents have used the Internet in the past 3 months, 89% use the Internet several times a day, and 96% access the Internet via a mobile phone [1]. Less is known about younger adolescents aged between 11 and 15 years. These younger adolescents fall into the category of “digital natives.” It represents a generation that has lived in an age where the Internet has always been around [2]. Yet, there is an expectation that adolescents use technology in everyday life while at their homes, schools, and during leisure time, it is surprising how

little there is research in the area of digital devices and apps for physical activity (PA).

Adolescence is an important period in a person's life. In addition to the biological changes to the structure of the body, in developed countries, many young adolescents start to reduce their reliance on their family and start to become independent and engage in social activities [3]. It is also a time of life where PA behaviors start to decline, or if sustained, seem to continue into adulthood [4]. Therefore, much attention has been directed to public health officials to act and provide appropriate contexts to reduce the drop off in PA levels [5,6].

The World Health Organization has provided a set of PA recommendations for people to meet, in order to benefit their health. These recommendations include taking part in at least 60 min of moderate to vigorous physical activities every day for children aged between 5 and 17 years [7]. In Finland, the levels of PA have been moderate (boys: 29.7%; girls: 17.7%) when compared with other European countries (boys: 23.4%; girls: 13.9%) [8]. For example, the proportion of 11-year old Finnish adolescents who meet the international PA recommendations [7] is the highest in Europe and North America [9]. However, the overall grade for various PA determinants in the Global matrix was a C, which is comparable with the European average [6]. Differences have been reported by gender, whereby more boys (23.4%) meet the PA recommendations than girls (13.9%) [8]. There is also a decline in PA levels as age increases from 49% in primary school to 18% in lower secondary school that meet the PA recommendations [10]. In addition, despite Finnish adolescents experience the least amount of inequality to other countries in Europe [11], there is an association between higher financial wealth status and more days of reported PA [12].

Commercial PA trackers are more popular than ever. The multifunctional purposes that these trackers possess include data about personal health and fitness management, which has made them into one of the top desirable digital consumer products [13]. Competition in the design and production is intense, with regular consumer and scientifically led tests between products [14]. Recent test results of commercial products have demonstrated some level of acceptability as well as room for improvement when reporting overall PA levels [15-18]. These devices can come in the form of a wristwatch, a mobile phone app as well as a noninvasive piece of equipment, such as a strap that monitors the heart rate [19]. There is also a much debate about the potential sensors have on personalized health [20]. Self-quantification is a growing area as it has the potential to inform individuals [21]. In addition, it has the capability to share information publicly or anonymously among peers. However, many of these studies neglect investigations from adolescents. Of studies that have investigated social media use by adolescents, many have been on risk behaviors [22], whereas a few actually explore the use in a sporting context [23].

Mobile phone apps are tools on the users' smartphone, whereas heart rate monitors or smart watches (HRMs/SWs) can operate independently from the phone. There are different levels of specifications from the hardware devices of HRM/SW. App user experiences are often richer than hardware devices, although there can be some overlap. For example, Polar HRM/SW has the capability to be linked with a mobile phone app, consequently the user has access to more data that the HRM/SW itself.

The majority of studies that used apps as a technique to positively change health behaviors were successful [24]. These findings are useful in the development of behavioral theories that connect personalized data with tracked behaviors, including PA, onto the Internet. People with health conditions or functional limitations, which are severe enough to cause a disability, have also benefited from apps [25,26]. Moreover, the apps are a popular tool used by adults in the mainstream population [27]. From this, there has been an assumption that PA trackers can support PA behaviors [28]. However, what many studies have done is to report the reliability of the devices [14,15,18], the benefits from wearing devices [28], as well as devices failing to have an effect on weight loss [29]. Intervention studies on adolescents have been rare [30], although the use of devices seems promising in its effect in reduction of screen-time, the long-term impact of such results are still unknown [31]. Ridgers and colleagues [30] found three interventions and two feasibility studies up until August 2016. In one of the intervention studies, Fitbit Ones were used a tool to promote PA to children with leukemia. However, despite overall happiness to use the PA tracker, there were no significant increases in step-count [32]. In comparison to studies with adults, the studies do not inform us whether the low number of studies reported is because of low usage among adolescents. To address this concern, the purpose of this study was to investigate the prevalence of adolescents who report on ownership and use PA measuring devices. In addition, this study examines the strength of association between ownership and use of devices and meeting adolescents PA recommendations.

Methods

Recruitment

The Finnish school-aged physical activity (SPA) study is a national representative cross-sectional study of many determinants of PA among adolescents aged between 11 and 15 years. Participants were selected at random, based on a stratified sampling method set by regional stratification in the capital, southern, central, and northern Finland. The class was the primary sampling unit through probability proportion size. In total, 109 Finnish-speaking schools and 65 Swedish-speaking schools participated. The response rates were 61% from the Finnish speaking schools, and 58% from the Swedish speaking schools [33]. Through proportional to size of school, the units were selected and a reserve list created to be used if necessary. The questionnaire was completed online in a computer class for a maximum of 60 min with an instructed teacher who presided over the data collection. The questionnaire was completed anonymously, voluntarily, and the study has been approved by the Institutional Ethical Review Board Committee.

Measures

Adolescents were asked to report their gender (boy or girl). They were also asked to report their month and year of birth.

The ages were then grouped into 11-, 13-, and 15-year old age groups based on their age closest to an age group category at the time of completing the survey.

PA Trackers

Ownership and usage of PA tracking devices were measured through two items. Both items began with the following opening question, "Do you have any of the following PA measuring devices...," and were followed by (1) app and (2) HRM/SW. The response categories of these measures were "I do not have," "Yes, but do not use it actively," and "Yes and use it actively." The wording used in the items were selected after extracting marketing materials of leading commercial PA measuring device companies in the Finnish language. The items were later translated into English for reporting.

Physical Activity Levels

Overall PA was measured through a single-item self-reported measure. As part of this measure, there was a pretext description of moderate to vigorous physical activity (MVPA). In addition, some example activities typical for adolescents to do were stated so that they could understand what to include in their recalled response [34]. Following this description, the respondents were asked, "Over the past 7 days, on how many days were you physically active for a total of at least 60 min per day?" Only one response was allowed that ranged from 0 to 7 days. For the study's purposes, the mean number of days was used to examine the differences in the average per group. In addition, a dichotomous cutoff based on the international PA recommendations [7] (7 days) versus not meeting the recommendations (between 0 and 6 days) was used for inferential statistics. The question has been widely used with acceptable validity and reliability in adolescent surveys [35,36]. The question has a significant validation correlation coefficient of .40 from adolescents in clinical settings [34] and is used widely in national and international adolescent health studies [37]. In Finland, the ICC values from a 2-week test-retest were between .7 and .8 [35]. Although there is criticism on the use of subjective reporting of PA, the use of self-reported PA in surveillance studies has been considered an approved method [38].

Disability

Data were disaggregated by disability through items used in the WHO Model Disability Survey [39]. Individuals were asked to report their severity of functional difficulties, including seeing, speaking, hearing, moving, breathing, and remembering or concentrating. These items were modified from the WHO Model Disability Survey to disaggregate by disability with a 5-point scale to indicate severity [40]. Severity was determined by functional difficulties that were of 1 (no difficulty), 2 (mild difficulty and did not affect participation), 3 (moderate difficulty and affects participation), 4 (severe difficult and affects participation), and 5 (complete difficulty and affects participation). To meet the criteria of having a disability, adolescents had to have reported at least one functional difficulty with severity of 3 or higher. In other words, the individual felt the difficulty affected their participation.

Family Affluence

The Family Affluence Scale (FAS) III is an international measure of social economic status of adolescents [41]. It comprises of six items related to family cars, holidays, computers, rooms, bathrooms, and dishwasher at home. An internationally established formula [41] was used to define three groups of family affluence: low, middle, and high [42].

Statistical Analysis

Study characteristics, such as age, gender, disability, and FAS were described through ownership and usage of mobile phone apps or HRMs/SWs, and were tested by chi-square test of independence ($P<.05$). Mean number of PA days were reported by each participant characteristic. Binary logistic regression analyst was conducted on the overall sample through dichotomized groups of participants that reported 0-6 days of MVPA and 7 days of MVPA. Adjusted odds ratios (OR) and 95% CIs were reported to indicate the likelihood of daily MVPA. To report positive odds, girls, who were aged 15 years, with low FAS and disability were reference groups. Furthermore, adolescents who had no PA trackers were also the reference groups. SPSS version 24.0 for Windows (SPSS Inc) was used for the analyses.

Results

Descriptive Statistics

In this study, half (2434/4575; 53.20%) of the 4575 adolescents (mean age 13.8 years, SD 1.64) were girls. Just over a quarter (1262/4575; 27.58%) of adolescents reported daily MVPA. Adolescents identified to have disabilities had indicated that their difficulty severity level was between "moderate" to "complete." Whereas, responses of severity levels of "mild" and "none" were grouped into adolescents without disabilities. On the basis of this definition, one in six (672/4568; 14.71%) adolescents reported to have disabilities. Family affluence was more skewed toward middle (2445/4575; 53.44%) and high (1650/4575; 36.07%) FAS, hence adolescents with low FAS (480/4575; 10.49%) were in the minority (Table 1).

Table 1. Sample characteristics of mean days of moderate to vigorous physical activity (MVPA) with 95% CI and proportion of meeting physical activity (PA) recommendations.

Characteristics	n	Mean (SD[a])	95% CI	% MVPA[b]
Gender				
Girls	2434	4.71 (1.82)	4.64-4.79	22.51
Boys	2141	5.05 (1.90)	4.96-5.12	33.35
Age groups (in years)				
11	1584	5.41 (1.77)	5.41-5.32	40.15
13	1594	4.90 (1.77)	4.81-4.99	25.72
15	1397	4.22 (1.89)	4.12-4.33	15.46
Disability				
With	672	4.44 (2.04)	4.28-4.59	23.21
Without	3896	4.94 (1.82)	4.88-5.00	28.29
Family affluence				
Low	480	4.64 (1.98)	4.46-4.82	25.63
Middle	2445	4.77 (1.90)	4.69-4.84	25.44
High	1650	5.07 (1.80)	4.99-5.16	31.33
HRM/SW[c]				
No	3223	4.72 (1.90)	4.65-4.78	25.04
Unused	782	5.06 (1.74)	4.94-5.18	28.77
Uses	408	5.70 (1.59)	5.54-5.86	44.12
Apps				
No	2116	4.65 (2.00)	4.57-4.73	25.09
Unused	1631	4.93 (1.78)	4.84-5.02	26.92
Uses	720	5.37 (1.67)	5.24-5.49	35.70

[a]SD: standard deviation.

[b]%MVPA: percentage that meet international physical activity recommendations for adolescents.

[c]HRM/SW: heart rate monitor or sport watch.

Of the adolescents who report ownership but not active use of apps, only (104/1579) 6.59% reported to use HRM/SW. Under one-third of app users (207/694; 29.8%) reported to use HRM/SW. Just over half (207/391; 52.9%) of HRM/SW users reported that they were users of apps too.

Mobile Phone Apps

More girls (1307/2351; 55.59%) reported that they have apps that measure PA than boys (1044/2351; 44.41%, *P*=.003). In addition, as the age of adolescents increased, the more apps owned to measure PA (*P*<.001). Among all the adolescents, one in six (720/4467; 16.12%) reported that they were using the apps to measure PA (Table 2).

Ownership of apps and meeting the PA recommendations was not always positive (OR =1.11, 95% CI 0.94-1.30). Apps usage was positively associated with the PA recommendations (OR 1.41, 95% CI 1.15-1.74) than adolescents without apps to track PA, after controlling for gender, age, disability, and FASIII.

Hear Rate Monitors and Sport Watches

More boys (590/2051; 28.77%) reported to own or use HRM/SW than girls 600/2362; 25.40%). One in ten (408/4413; 9.25%) adolescents reported to use HRM/SW. Almost half (325/694; 46.8%) of the active users of mobile phone apps that measure PA reported to have HRM/SW, although only 29.8% (207/694) were also using HRM/SW. There were more ownership and usage of HRM/SW in adolescents with higher family affluence groups (Table 3).

There were stronger associations with HRM/SW and PA. The association between owning HRM/SW and meeting the PA recommendations was not necessarily positive all the time (OR 1.18, 95% CI 0.97-1.42). Adolescents who owned and used HRM/SW were 2.09 times (95% CI 1.64-2.67) more likely to meet the PA recommendations than adolescents without HRM/SW, after controlling for gender, age, disability, and FAS III.

Table 2. Sample characteristics of users of smartphone apps and chi-square test of independence.

Characteristics	Smartphone apps				*P* value
	Total, n	None (%)	Yes and unused (%)	Yes and used (%)	
Gender					.01
Boys	2077	1033 (49.74)	717 (34.52)	327 (15.74)	
Girls	2390	1083 (45.31)	914 (38.24)	393 (16.44)	
Age groups (in years)					<.001
11	1537	788 (51.27)	490 (31.88)	259 (16.85)	
13	1560	721 (46.21)	591 (37.88)	248 (15.90)	
15	1370	607 (44.31)	550 (40.15)	248 (18.10)	
Disability					.75
With	652	313 (48.0)	230 (35.3)	109 (16.7)	
Without	3810	1802 (47.29)	1400 (36.75)	608 (15.96)	
Family affluence					<.001
Low	474	291 (61.4)	123 (25.9)	60 (12.7)	
Middle	2381	1199 (50.35)	856 (35.95)	326 (13.69)	
High	1612	626 (38.83)	652 (40.45)	334 (20.72)	
60 min MVPA[a]					<.001
Not daily	3240	1585 (48.92)	1192 (36.79)	463 (14.29)	
Daily	1227	531 (43.28)	439 (35.78)	257 (20.95)	
Total	4467	2116 (47.37)	1631 (36.51)	720 (16.12)	

[a]MVPA: moderate to vigorous physical activity.

Table 3. Sample characteristics for users of heart rate monitors or smart watches and chi-square test of independence.

Characteristics	Heart rate monitor or sports watch				*P* value
	Total, n	None (%)	Yes unused (%)	Yes and used (%)	
Gender					.04
Boys	2051	1461 (71.23)	387 (18.87)	203 (9.90)	
Girls	2362	1762 (74.60)	395 (16.72)	205 (8.68)	
Age groups (in years)					.07
11	1517	1140 (75.15)	235 (15.49)	142 (9.36)	
13	1542	1113 (72.18)	284 (18.42)	145 (9.40)	
15	1354	970 (71.64)	263 (19.42)	121 (8.94)	
Disability					.52
With	637	461 (72.4)	122 (19.2)	54 (8.5)	
Without	3770	2761 (73.24)	659 (17.48)	350 (9.28)	
Family affluence					<.001
Low	466	383 (82.2)	62 (13.3)	21 (4.5)	
Middle	2357	1814 (76.96)	357 (15.15)	186 (7.89)	
High	1590	1026 (64.53)	363 (22.83)	201 (12.64)	
60 min MVPA[a]					<.001
Not daily	3201	2416 (75.48)	557 (17.40)	228 (7.12)	
Daily	1212	807 (66.58)	225 (18.56)	180 (14.85)	
Total	4413	3233 (73.26)	782 (17.72)	408 (9.25)	

[a]MVPA: moderate to vigorous physical activity.

Discussion

Principal Findings

The findings of this study suggest that over half of adolescents own mobile phone apps that can function as PA trackers, however, one in six reported to actually use the apps. In addition, at the time of the study, a quarter of adolescents own HRM/SW yet, one in ten used HRM/SW.

Phone Apps Are Popular for PA Tracking

In light of the official statistics in Finland, by the time adolescents reach the age of 16 years, almost all (96%) of adolescents can access the Internet via mobile phones [1]. This high usage that was expected with the main reason has attributed to the absence of landlines in Finland since the 1990s and has encouraged mobile phones use among adolescents [43]. The findings from this study reveal the majority of adolescents (2351/4467; 52.63%) have sufficient interest in PA behaviors to have PA tracking apps on their phones.

A poll found that downloading apps were the third most popular function (82%), behind taking photos and messaging friends in adolescents aged between 13 and 16 years from the United Kingdom [44]. Data on electronic media usage reveals that three quarters of Finnish adolescents aged between 10 and 14 years use mobile phones and social media takes up the majority of their free time [45]. Other activities such as social media, surfing the Internet, playing games, watching videos, reading, and listening to music can take up the most amount of time on electronic media. Not all smart phones have, by default, PA tracking apps. However, they can include hardware components, such as accelerometers and GPS, which are key functions to track PA. To enable these hardware components, users often need to be aware of the functions, activate the apps, and feel the desire to install the apps.

Previous research has shown PA as a socially desirable habit [46], although in the case of these adolescents, one in six reported to use the apps for the purposes of monitoring PA. The availability of these apps on the phones, and low usage highlights the context in which adolescents use the mobile phone as an organic part of everyday life [43]. App designers may need to target adolescents who carry their phone with them everywhere, by encouraging a fun way that is also purposeful to meet their daily electronic usage needs and at the same time for self-input PA data. Games like Pokémon Go seem to have a mechanism to achieve this. However, an update to this would be providing more support by tailoring their existing PA levels and providing feedback on habitual PA. This may spur on interest in the differences between simply owning an app and actual usage of apps.

There was a positive association between the usage of apps and meeting the current international PA recommendations. The apps connect with various sensors on the phone and translate the data from algorithms to the user interface [47]. Despite the varying levels of accuracy when the apps produce information [24,48], it is a starting point for quantifying the self in PA behaviors [21]. This type of self-quantification needs to be

explored among digital natives since previous research has excluded responses by adolescents aged between 14 and 17 years [13].

The international PA recommendations are there to encourage children aged between 5 and 17 years to take part in daily MVPA [7]. According to the results from this study, the odds of active users of mobile phone apps to meet the recommendations were 1.4 compared with the reference group of nonowners of PA tracking apps. Apps have enormous potential to provide personal and immediate behavioral feedback. In control theory [49], feedback loops can direct behavior toward desirable patterns. The positive reinforcement by the feedback is more than self-quantification. Adolescents can use the apps and connect to other users [47]. The individual can share their results with their friends. This can encourage users to be more physically active, to demonstrate the recommendations are met, as well as bring up a point of discussion between users for ways to take part in enjoyable activities.

Apps that display information for personalized feedback can motivate individuals to be more physically active [50]. This is an important health promotion opportunity. One of the main ways to combat noncommunicable disease is to decrease physical inactivity. Along with the popular behavior theories to support behavioral change techniques [51], researchers may need to start thinking about how apps play a role in the advancement of such developments. The results from this study highlight the need for developers of apps and researchers who use apps for interventions to recognize that an app is not enough. Often one-size-fits-all in traditional health promoting interventions do not work [52]. Through changes in software, apps that read individual data can give more personalized feedback [30]. This is important to reach the populations that are physically inactive since, setting sights on an immediate change to daily MVPA may seem too off putting for them.

Compared with other intervention methods, one main advantage that apps have over other techniques are the tailoring mechanisms for each individual. Apps can be programmed to create easy-to-reach goals in the early stages of PA promotion, and then slowly progress toward the PA recommendations and beyond with the user in control. Therefore, it is recommended that there are apps developments that take into account segmentation in gender, age, dis/ability, and wealth.

Heart Rate Monitors or Sports Watches as Specific Devices

Devices such as HRM/SW require extra user commitment such as maintenance, care, battery charging, turning on and off. Therefore, it was no surprising that users of HRM/SW had the largest ORs for meeting PA recommendations. Moreover, it is most likely that users of HRM/SW are early adopters of PA tracking devices as they have already invested a large amount of their time in PA and sports. The feedback loop, under control theory [49], from HRM/SW has more emphasis than in apps. Feedback from HRM/SW is readily available from a display, usually positioned on the wrist. A quick glance can modify motivation levels to take part in more activity so that the individual meets the daily targets. Alternatively, the user may be looking for confirmation that they have reached their target so that they can choose other activities, such as school work, socializing, arts, or just relaxing. These prompts are inviting for users who would want a device to help them meet the PA recommendations. However, motivational nudges to change behavior have the risk of polarization. Whereby individuals who do not meet the recommendations due to other difficulties, resist the feedback and find new gadgets distasteful. Feasibility studies from adolescents suggest that comfort, design, and feedback features are more important than the usability and effectiveness for increasing PA behaviors [30]. As with apps, interventions with HRM/SW may need a tailoring approach for effective health promotion in adolescents.

Today's PA devices worn on the wrist have the capability to measure the number of steps taken during the day and can be compared with the recommended step count level that is equivalent to 60 min of MVPA [53,54]. Physical and health educators may want to start considering how to integrate this knowledge into their messages. Furthermore, should HRM/SW be proven to be effective in PA promotion, national agencies may want to consider the creation of subsides to assist purchasing of equipment, thus allowing pupils to learn about self-quantification in the classes. This development is only the beginning, as newer devices become more readily available, with more functions, more feedback preferences, and at lower costs [48].

Subpopulation Differences in Ownership and Usage of PA Trackers

Tailoring of PA trackers may also need to consider the gender differences reported in this study. More girls reported that they had apps to measure PA than boys. Goal orientation toward PA tends to differ between boys and girls; for example, girls tend to take part in PA for enjoyment, whereas boys take part in PA for its competitive nature [55]. Another issue could be the changes in maturation between the genders during these ages. During early adolescents, studies have reported that girls have stronger digital literacy skills than boys [56]. Boys may know less about the use of apps on the mobile phones [57]. However, more boys who reported to have PA trackers reported to actually use the apps and HRM/SW. The mechanisms to use the trackers may be common only among boys who have an interest in the behavior. Consequently, this would indicate that the use of PA trackers were to support competitive motives through PA rather than to be used as an intervention tool to promote more PA behaviors.

There was an increase in the number of users of PA trackers as the cohorts increased with age. Traditionally, there have been substantial declines in PA levels to the users who are aged between 11 and 15 [12]. Many potential benefits from these designs are possible. At the age of 11 years, adolescents are usually active, requiring fewer interventions than 15-year-olds. Usually at the turn from young to mid adolescents, there is a need to fit in socially [3]. Technology use to aid social interactions is second nature for digital natives. Understanding this phenomenon by digital natives is an important consideration for health promoters as some traditional methods have low relevance for young adolescents.

It was no surprise that high family affluence was positively associated with PA trackers. Although the majority of apps are freely available to download from the mobile phone app stores, families must provide adolescents with phones that have sufficient capabilities for measuring PA. Moreover, the combination of peripheral devices (HRM/SW) can be quite an expensive acquisition for some families. Previous studies have reported the association with higher family affluence and increases in PA levels in adolescents [12], and it now extenuated with ownership of HRM/SW, which was also associated with more days of reported PA.

Adolescents without disabilities reported more days of MVPA than adolescents with disabilities. However, the differences were not strong enough to influence the statistical model. This concurs with previous studies from mainstream school populations whereby there were no significant differences in PA levels between adolescents with and without disabilities [58,59]. In this study, the differences among the various functional differences were not reported neither since, of all the different functional difficulty categories, only adolescents with moving difficulties were significantly less active than other adolescents [60]. The usage and ownership of PA trackers were also indifferent between adolescents with and without disabilities. These results can be used to inform health practitioners to use mobile health apps and other devices as part of targeted interventions for adolescents despite their health conditions.

Study Limitations

The results of this study have some limitations. Data collected were from a cross-sectional study design. Therefore, the associations between devices and PA are not causal. In other words, the results should equally be read so that adolescents who are more interested in being physically active have devices, as adolescents have devices do so to taking part in more daily PA. Intervention and longitudinal studies might provide better insight into the causal links, behavioral techniques to increase PA levels, and commercialization of products to reinforce the sense of athletic identity. More studies that utilize longitudinal designs to investigate the relationships between PA trackers and its influences on PA are strongly recommended. PA levels were collected through self-report methods. There may be some over or under reporting bias; however, because all respondents used the same method, it has been possible to find the associations between PA trackers and PA levels. In addition, the items that measured PA trackers have not been tested for meeting minimum reliability and validity scores. The results of the studies were limited to an undefined use of apps or HRM/SW. Future studies may consider development of these items to help differentiate the types of users across activity trackers. Future studies may need to improve the items used to eliminate as much error as possible.

Conclusions

This study has reported the ownership and usage of PA trackers among adolescents in Finland. It is an exciting time for researchers as digital natives use accessible tools for their own personal health. In this study, the appeal of PA trackers was evident, as over half of Finnish adolescents have mobile phone apps to track PA. The strongest association reported was among the one in ten adolescents that have HRM/SW and took part in daily PA to correspond with the current PA recommendations for health [7]. However, only 22.51% (548/2434) and 33.35% (714/2141) of girls and boys, respectively, reported to meet these recommendations. Other differences in age and wealth were observed, although not in disability. More understanding is needed on the mechanisms to build suitable apps and devices for adolescents as they live through bodily and supportive changes in the life, while promoting healthy behaviors. Apps and device development would benefit from understanding the individual differences in current PA levels and relative targets.

Acknowledgments

The authors thank Jari Villberg for help with the data request and Riikka Hämylä for help with data collection. This study was funded by the Ministry of Education and Culture (grant number: OKM/89/626/2015).

Conflicts of Interest

None declared.

References

1. Official Statistics of Finland. Use of informationcommunications technology by individuals URL: http://www.stat.fi/meta/til/sutivi_en.html [accessed 2017-04-14] [WebCite Cache ID 6piW3L5TV]
2. Underwood J, Farrington-Flint L. Learning and the E-generation. Chichester: John Wiley & Sons; 2015.
3. Sawyer SM, Drew S, Yeo MS, Britto MT. Adolescents with a chronic condition: challenges living, challenges treating. The Lancet 2007 Apr;369(9571):1481-1489. [doi: 10.1016/S0140-6736(07)60370-5]
4. Wichstrøm L, von ST, Kvalem IL. Predictors of growth and decline in leisure time physical activity from adolescence to adulthood. Health Psychol 2013 Jul;32(7):775-784. [doi: 10.1037/a0029465] [Medline: 22924445]
5. Rütten A, Schow D, Breda J, Galea G, Kahlmeier S, Oppert J, et al. Three types of scientific evidence to inform physical activity policy: results from a comparative scoping review. Int J Public Health 2016 Jun;61(5):553-563 [FREE Full text] [doi: 10.1007/s00038-016-0807-y] [Medline: 27113707]

6. Tremblay M, Barnes J, González SA, Katzmarzyk PT, Onywera VO, Reilly JJ, Global Matrix 2.0 Research Team. Global Matrix 2.0: report card grades on the physical activity of children and youth comparing 38 countries. J Phys Act Health 2016 Nov;13(11 Suppl 2):S343-S366. [doi: 10.1123/jpah.2016-0594] [Medline: 27848745]
7. World Health Organization. Global recommendations on physical activity for health. Geneva, Switzerland: World Health Organization; 2010.
8. Kalman M, Inchley J, Sigmundova D. Secular trends in moderate-to-vigorous physical activity in 32 countries from 2002 to 2010: a cross-national perspective. Eur J Public Health 2015;25(Suppl 2):37-40. [doi: 10.1093/eurpub/ckv024]
9. Inchley J, Currie D, Young T. Growing up unequal: gender and socioeconomic differences in young people's health and well-being. In: Health Behaviour in School-aged Children (HBSC) Study: International Report from the 2013/2014 Survey. Copenhagen: WHO Regional Office for Europe; 2016.
10. Tammelin TH, Aira A, Hakamäki M, Husu P, Kallio J, Kokko S, et al. Results from Finland's 2016 report card on physical activity for children and youth. J Phys Act Health 2016 Nov;13(11 Suppl 2):S157-S164. [doi: 10.1123/jpah.2016-0297] [Medline: 27848744]
11. Chzhen Y, Bruckauf Z, Ng K, Pavlova D, Torsheim T, de Matos M. UNICEF. Geneva: UNICEF; 2016. Inequalities in adolescent health and life satisfaction: evidence from the health behaviour in school-aged children study URL: https://www.unicef-irc.org/publications/pdf/IWP_2016_09.pdf [accessed 2017-04-26] [WebCite Cache ID 6q0SkYbB1]
12. Borraccino A, Lemma P, Iannotti RJ, Zambon A, Dalmasso P, Lazzeri G, et al. Socioeconomic effects on meeting physical activity guidelines. Med Sci Sports Exerc 2009;41(4):749-756. [doi: 10.1249/mss.0b013e3181917722] [Medline: 19276860]
13. Accenture. 2014. Racing toward a complete digital lifestyle: digital consumers crave more URL: https://www.accenture.com/t20150826T034315__w__/pl-en/_acnmedia/Accenture/Conversion-Assets/Careers/Documents/LandingPage/1/Accenture-Digital-Consumer-Tech-Survey-2014.pdf [accessed 2017-03-02] [WebCite Cache ID 6oeiv2diY]
14. Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de GM. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 2015 Oct;7:24 [FREE Full text] [doi: 10.1186/s13102-015-0018-5] [Medline: 26464801]
15. Lee J, Kim Y, Welk GJ. Validity of consumer-based physical activity monitors. Med Sci Sports Exerc 2014 Sep;46(9):1840-1848. [doi: 10.1249/MSS.0000000000000287] [Medline: 24777201]
16. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. J Sci Med Sport 2014 Sep;17(5):496-500. [doi: 10.1016/j.jsams.2013.10.241] [Medline: 24268570]
17. 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;7:952 [FREE Full text] [doi: 10.1186/1756-0500-7-952] [Medline: 25539733]
18. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA 2015 Feb 10;313(6):625-626. [doi: 10.1001/jama.2014.17841] [Medline: 25668268]
19. 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]
20. Appelboom G, Camacho E, Abraham ME, Bruce SS, Dumont EL, Zacharia BE, et al. Smart wearable body sensors for patient self-assessment and monitoring. Arch Public Health 2014 Aug;72(1):28 [FREE Full text] [doi: 10.1186/2049-3258-72-28] [Medline: 25232478]
21. Almalki M, Gray K, Martin-Sanchez F. Activity theory as a theoretical framework for health self-quantification: a systematic review of empirical studies. J Med Internet Res 2016 May 27;18(5):e131. [doi: 10.2196/jmir.5000]
22. Yonker LM, Zan S, Scirica CV, Jethwani K, Kinane TB. "Friending" Teens: systematic review of social media in adolescent and young adult health care. J Med Internet Res 2015 Jan 05;17(1):e4. [doi: 10.2196/jmir.3692]
23. Lauritsalo K, Saakslahti A, Rasku-Puttonen H. Student'svoice online: experiences of PE in Finnish schools. APE 2012;02(03):126-131. [doi: 10.4236/ape.2012.23022]
24. 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]
25. van den Berg-Emons RJ, Bussmann JB, Stam HJ. Accelerometry-based activity spectrum in persons with chronic physical conditions. Arch Phys Med Rehabil 2010 Dec;91(12):1856-1861. [doi: 10.1016/j.apmr.2010.08.018] [Medline: 21112426]
26. Dobkin BH, Dorsch A. The promise of mHealth: daily activity monitoring and outcome assessments by wearable sensors. Neurorehabil Neural Repair 2011;25(9):788-798 [FREE Full text] [doi: 10.1177/1545968311425908] [Medline: 21989632]
27. 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;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
28. Sanders JP, Loveday A, Pearson N, Edwardson C, Yates T, Biddle SJ, et al. Devices for self-monitoring sedentary time or physical activity: a scoping review. J Med Internet Res 2016 May 04;18(5):e90 [FREE Full text] [doi: 10.2196/jmir.5373] [Medline: 27145905]
29. 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]

30. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. JMIR Mhealth Uhealth 2016 Nov 23;4(4):e129 [FREE Full text] [doi: 10.2196/mhealth.6540] [Medline: 27881359]
31. Lubans DR, Smith JJ, Peralta LR, Plotnikoff RC, Okely AD, Salmon J, et al. A school-based intervention incorporating smartphone technology to improve health-related fitness among adolescents: rationale and study protocol for the NEAT and ATLAS 2.0 cluster randomised controlled trial and dissemination study. BMJ Open 2016 Jun 27;6(6):e010448 [FREE Full text] [doi: 10.1136/bmjopen-2015-010448] [Medline: 27354070]
32. Hooke MC, Gilchrist L, Tanner L, Hart N, Withycombe JS. Use of a fitness tracker to promote physical activity in children with acute lymphoblastic leukemia. Pediatr Blood Cancer 2016 Apr;63(4):684-689. [doi: 10.1002/pbc.25860] [Medline: 26756736]
33. JYU. Helsinki: National Sports Council; 2016. The physical activity behaviours of children and adolescents in Finland; results of the LIITU study. [Lasten ja Nuorten liikuntakäyttäytyminen Suomessa, LIITU-tutkimuksen tuloksia 2016] URL: https://www.jyu.fi/sport/laitokset/tutkimusyksikot/tetk/liitu [accessed 2017-04-26] [WebCite Cache ID 6q0TgE1nb]
34. Prochaska JJ, Sallis JF, Long B. A physical activity screening measure for use with adolescents in primary care. Arch Pediatr Adolesc Med 2001 May 01;155(5):554. [doi: 10.1001/archpedi.155.5.554]
35. Vuori M, Ojala S, Tynjälä J. The stability of questions concerning physical activity in the HBSC study. [Liikunta-aktiivisuutta koskevien kysymysten stabiliteetti WHO-koululaistutkimuksessa]. Liikunta & Tiede 2005;42(6):39-46.
36. Bobakova D, Hamrik Z, Badura P, Sigmundova D, Nalecz H, Kalman M. Test-retest reliability of selected physical activity and sedentary behaviour HBSC items in the Czech Republic, Slovakia and Poland. Int J Public Health 2015 Jan;60(1):59-67. [doi: 10.1007/s00038-014-0628-9] [Medline: 25471078]
37. Sallis JF, Bull F, Guthold R, Heath GW, Inoue S, Kelly P, et al. Progress in physical activity over the Olympic quadrennium. The Lancet 2016 Sep;388(10051):1325-1336. [doi: 10.1016/S0140-6736(16)30581-5]
38. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. The Lancet 2012 Jul;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1]
39. UN. 2015. Disability data and statistics, monitoring and evaluation: the way forward- a disability-inclusive agenda towards 2015 and beyond URL: http://www.un.org/disabilities/documents/egm2014/EGM_FINAL_08102014.pdf [accessed 2017-03-02] [WebCite Cache ID 6oejwTg9a]
40. Tardi R, Njelesani J. Disability and the post-2015 development agenda. Disabil Rehabil 2015 Oct;37(16):1496-1500. [doi: 10.3109/09638288.2014.972589] [Medline: 25350660]
41. Hartley J, Levin K, Currie C. A new version of the HBSC family affluence scale - FAS IIIcottish qualitative findings from the International FAS development study. Child Indic Res 2015;9(1):233-245. [doi: 10.1007/s12187-015-9325-3]
42. Torsheim T, Cavallo F, Levin KA, Schnohr C, Mazur J, Niclasen B, FAS Development Study Group. Psychometric validation of the revised family affluence scale: a latent variable approach. Child Indic Res 2016 Oct;9:771-784 [FREE Full text] [doi: 10.1007/s12187-015-9339-x] [Medline: 27489572]
43. Oksman V, Turtiainen J. Mobile communication as a social stage: meanings of mobile communication in everyday life among teenagers in Finland. New Media & Soc 2004 Jun 01;6(3):319-339. [doi: 10.1177/1461444804042518]
44. Statista. Smartphone usage behavior among children in Great Britain 2014, by age group URL: https://www.statista.com/statistics/293508/smartphone-usage-behavior-among-children-by-age-great-britain/ [accessed 2017-03-02] [WebCite Cache ID 6oehjwMaf]
45. Merikivi J, Myllyniemi S, Salasuo M. Media glove. Children's and young people's leisure-time study in 2016 on media and sport. [Media Hanskassa; Lasten ja nuorten vapaa-aikatutkimus]. Helsinki: Finnish Youth Research Society; 2016.
46. Brener ND, Billy JO, Grady WR. Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: evidence from the scientific literature. J Adolesc Health 2003 Dec;33(6):436-457. [doi: 10.1016/S1054-139X(03)00052-1]
47. Lowe SA, Ólaighin G. Monitoring human health behaviour in one's living environment: a technological review. Med Eng Phys 2014 Feb;36(2):147-168. [doi: 10.1016/j.medengphy.2013.11.010] [Medline: 24388101]
48. Johnston W, Hoffman S, Thornton L. Mobile health: a synopsis and comment on “Increasing physical activity with mobile devices: a meta-analysis”. Transl Behav Med 2014 Mar;4(1):4-6 [FREE Full text] [doi: 10.1007/s13142-014-0254-3] [Medline: 24653771]
49. Carver CS, Scheier MF. Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. Psychol Bull 1982;92(1):111-135. [doi: 10.1037/0033-2909.92.1.111]
50. 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;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
51. 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]
52. Wise J. NICE chief criticises “one size fits all” approach to health promotion. BMJ 2009 Jun 10;338(jun10 1):b2326-b2326. [doi: 10.1136/bmj.b2326]

53. Adam NJ, Spierer DK, Gu J, Bronner S. Comparison of steps and energy expenditure assessment in adults of Fitbit tracker and ultra to the actical and indirect calorimetry. J Med Eng Technol 2013 Oct;37(7):456-462. [doi: 10.3109/03091902.2013.831135] [Medline: 24007317]
54. Dontje ML, de GM, Lengton RR, van der Schans C, Krijnen WP. Measuring steps with the Fitbit activity tracker: an inter-device reliability study. J Med Eng Technol 2015 Jul;39(5):286-290. [doi: 10.3109/03091902.2015.1050125] [Medline: 26017748]
55. Biddle SJ, Atkin AJ, Cavill N, Foster C. Correlates of physical activity in youth: a review of quantitative systematic reviews. Int Rev Sport Exerc Psychol 2011 Mar;4(1):25-49. [doi: 10.1080/1750984X.2010.548528]
56. Garbe C, Lafontaine D, Shiel G. European Literacy Policy Network (ELINET). 2015. Literacy in Finland: Country Report - Children and Adolescents URL: http://www.eli-net.eu/fileadmin/ELINET/Redaktion/user_upload/Childrenadolescents_Long_Finland_final.pdf [accessed 2017-03-02] [WebCite Cache ID 6oejV4pcY]
57. Kupiainen R, Suoninen A, Nikunen K. Online habits of Finnish children; use, risks and data misuse. Nordicom-Information 2011;33(4):51-57.
58. Ng K, Rintala P, Tynjälä J, Välimaa R, Villberg J, Kokko S, et al. physical activity trends of Finnish adolescents with long-term illnesses or disabilities from 2002-2014. J Phys Act Health 2016 Aug;13(8):816-821. [doi: 10.1123/jpah.2015-0539] [Medline: 26999825]
59. Rintala P, Välimaa R, Tynjälä J, Boyce W, King M, Villberg J, et al. Physical activity of children with and without long-term illness or disability. J Phys Act Health 2011 Nov;8(8):1066-1073. [doi: 10.1123/jpah.8.8.1066]
60. Ng K, Rintala P, Saari A. Functional Difficulties, physical activity, sedentariness [Toimintakyvyn ja -rajoitteiden yhteydet liikunta-aktiivisuuteen ja paikallaanoloon]. In: Kokko S, Mehtälä A, editors. The Finnish School-aged Physical Activity Mointoring Study: Results from 2016 study [Lasten ja nuorten liikuntakäyttäytyminen Suomessa (LIITU) -tutkimuksen aineistonkeräys ja menetelmät 2016]. Helsinki: National Sports Council; 2016:73-78.

Abbreviations

FAS: Family Affluence Scale III
HRM/SW: heart rate monitor or sports watch
MVPA: moderate to vigorous physical activities
OR: odds ratio
PA: physical activities
SPA: school-aged physical activity
WHO: World Health Organization

Edited by G Eysenbach; submitted 13.12.16; peer-reviewed by C Hawke, C Shen, E Lyons, E Dooley; comments to author 03.02.17; revised version received 02.03.17; accepted 23.03.17; published 04.05.17.

Please cite as:
Ng K, Tynjälä J, Kokko S
Ownership and Use of Commercial Physical Activity Trackers Among Finnish Adolescents: Cross-Sectional Study
JMIR Mhealth Uhealth 2017;5(5):e61
URL: http://mhealth.jmir.org/2017/5/e61/
doi:10.2196/mhealth.6940
PMID:28473304

Original Paper

Assessing the Impact of a Novel Smartphone Application Compared With Standard Follow-Up on Mobility of Patients With Knee Osteoarthritis Following Treatment With Hylan G-F 20: A Randomized Controlled Trial

Nebojsa Skrepnik[1*], MD, PhD; Andrew Spitzer[2*], MD; Roy Altman[3*], MD; John Hoekstra[4], MD, PhD; John Stewart[5*], MSc; Richard Toselli[6*], MD

[1]Tucson Orthopaedic Institute, Tucson, AZ, United States
[2]Cedar-Sinai Orthopaedic Center, Los Angeles, CA, United States
[3]University of California Los Angeles Medical Center, Los Angeles, CA, United States
[4]National Clinical Research–Richmond, Richmond, VA, United States
[5]Sanofi, Laval, QC, Canada
[6]Sanofi, Paris, France
[*]these authors contributed equally

Corresponding Author:
Nebojsa Skrepnik, MD, PhD
Tucson Orthopaedic Institute
5301 E. Grant Road
Tucson, AZ,
United States
Phone: 1 520 784 6140
Fax: 1 520 784 6146
Email: NSkrepnik@tucsonortho.com

Abstract

Background: Osteoarthritis (OA) is a leading cause of disability in the United States. Although no disease-modifying therapies exist, patients with knee OA who increase walking may reduce risk of functional limitations.

Objective: The objective of the study is to evaluate the impact of a mobile app (OA GO) plus wearable activity monitor/pedometer (Jawbone UP 24) used for 90 days on the mobility of patients with knee OA treated with hylan G-F 20.

Methods: Patients with knee OA aged 30 to 80 years who were eligible to receive hylan G-F 20 and were familiar with smartphone technology were enrolled in this randomized, multicenter, open-label study. Patients who had a body mass index above 35 kg/m^2 were excluded. All patients received a single 6-mL injection of hylan G-F 20 and wore the Jawbone monitor. The patients were then randomized 1:1 to Jawbone and OA GO (Group A; n=107) with visible feedback (unblinded) or Jawbone only (Group B; n=104) with no visible feedback (blinded). The primary endpoint was mean change from baseline in steps per day at day 90 between Groups A and B.

Results: Baseline characteristics were similar between groups. There were significant differences between the increases in least squares (LS) mean number of steps per day (1199 vs 467, *P*=.03) and the mean percentage change (35.8% vs 11.5%, *P*=.02) from baseline in favor of Group A over Group B. There was a greater reduction in pain from baseline during the 6-minute walk test in Group A versus Group B. (LS mean change: −55.3 vs −33.8, *P*=.007). Most patients (65.4%) and surveys of physicians (67.3%) reported they would be likely or very likely to use/recommend the devices. Patient Activity Measure-13 scores improved from baseline (LS mean change for Groups A and B: 5.0 vs 6.9), with no significant differences between groups. The occurrence of adverse events was similar in the 2 groups.

Conclusions: Use of a novel smartphone app in conjunction with a wearable activity monitor provided additional improvement on mobility parameters such as steps per day and pain with walking in the 6-minute walk test in patients with knee OA who were treated with hylan G-F 20. Results also highlight the amenability of patients and physicians to using mobile health technology in the treatment of OA and suggest further study is warranted.

(JMIR Mhealth Uhealth 2017;5(5):e64) doi:10.2196/mhealth.7179

KEYWORDS

mobile health; mHealth; mobile apps; osteoarthritis; osteoarthritis, knee; hylan G-F 20; Synvisc

Introduction

Osteoarthritis

Osteoarthritis (OA) is a leading cause of disability in the United States [1]. From 2010 to 2012, an estimated 52.5 million (22.7%) US adults reported doctor-diagnosed arthritis, representing a net increase of 0.87 million adults with arthritis per year since the 2007 to 2009 estimate of 49.9 million [2]. The prevalence of OA, the most common form of arthritis [1], is estimated to double by 2020 [3] as predicted from the increase in the number of US adults with clinical hand, hip, or knee OA from 21 million in 1995 to 27 million in 2005 [4]. Estimates suggest that symptomatic knee OA occurs in 13% of women and 10% of men aged 60 years and older [5], approximately 17% of adults aged 45 years and older [4], and approximately 5% of adults aged 26 years and older [4]. Furthermore, the number of people affected by symptomatic OA is likely to increase as the population ages and the rate of obesity increases [5]. OA of the knee can significantly contribute to pain and lack of physical activity [2,6,7]. The resulting reduction in mobility has also been found to negatively impact quality of life [8] and result in an increased risk of sick leave and need for disability [9]. A study using US national survey data found that OA-associated absenteeism costs equaled approximately 3 lost work days per year, totaling up to $10.3 billion in annual absenteeism costs [10].

Management of Knee Osteoarthritis

To date, there is no disease-modifying therapy available for OA [11]. Optimal management of knee OA requires a combination of both pharmacologic (eg, oral or topical analgesics or intra-articular therapies [viscosupplementation, corticosteroids]) and nonpharmacologic methods (eg, weight loss, exercise, and physical activity) [12-15]. However, despite the well-recognized benefits of exercise and the promotion of physical activity by many professional societies [12,13,15], there is no standard exercise or education program and no clear benefit of one exercise program over another [13]. Information about the benefits of exercise is readily available [12,13,15] but rarely incorporated into patient behavior [16,17], and adherence to such programs by patients with knee OA is poor [16,17]. There is also no consensus on which measures or combination of measures should be used to assess physical function in patients with knee OA [18], complicating the initial assessment and tracking of progress in mobility improvement.

Walking has been shown to significantly reduce symptoms and the risk of functional limitations due to knee OA [19,20]. In one study, patients with OA of the knee who walked 6,000 steps per day or more reduced the risk of developing a functional limitation by half within the next 2 years [19]. This suggests that walking may maintain knee function. However, strategies to encourage adherence to regular walking are clearly needed.

Benefits of Mobile Health Apps

Adoption of mobile health apps has the potential to improve patient outcomes. Mobile health apps have been shown to be a useful tool in weight loss programs [21,22]. In addition, the use of short message service text message reminders has been shown to improve health care appointment attendance across health care settings [23], and there is evidence from a randomized, controlled trial to suggest that monthly phone contact can result in clinical improvements in patients with knee OA [24]. Patient physical activity level has been found to increase in studies using Internet- and mobile-based apps for postrehabilitation exercise persistence in chronic obstructive pulmonary disease [25], in physical activity maintenance following cardiac rehabilitation [26], in regular physical activity engagement by cancer survivors [27], in increasing walking based in the workplace [28], and in promoting a healthy lifestyle [29].

The use of mobile apps in health care may be one such strategy to increase physical activity and enhance OA management. However, there are currently no published data on the use of mobile apps for the management of knee OA.

The objective of this study was to examine the impact on mobility in knee OA patients who were treated with hylan G-F 20 and standard of care follow-up plus a mobile smartphone app (OA GO) and a wearable activity monitor (Jawbone UP 24) versus hylan G-F 20 and standard of care follow-up only over 90 days. We hypothesized that patients using the OA GO app and a wearable activity monitor would experience a positive impact on their mobility compared with those not using the app.

Methods

Study Population

Consecutive patients with knee OA whom the physician investigator decided to treat with one 6-mL injection of hylan G-F 20 (in accordance with the US label) who consented and qualified were enrolled in this open-label, multicenter, randomized, parallel-group study. Study sites and physicians were recruited using a detailed questionnaire and site qualification visit. Patients were recruited from the selected private community-based practices and research-only practice sites. All eligible patients were able to read and understand English and provided informed consent prior to starting the study. The protocol complied with the 18th World Health Congress (Helsinki, 1964) recommendations and applicable amendments and with any applicable country-specific laws, regulations, and guidelines. The protocol was approved by relevant ethics committees and/or institutional review boards. After working with the US Food and Drug Administration (FDA) to initiate registration of the trial, the sponsor deemed the study not eligible to be posted on ClinicalTrials.gov considering that clinical studies using devices whereby the primary outcome measure relates solely to feasibility and not to health outcomes are excluded from registration. The FDA

concurred there was not a requirement to register the trial at that time.

To be included in the study, all patients must have had unilateral knee OA and have been suitable for treatment with hylan G-F 20 based on the decision of the physician investigator. Study sites were requested to report whether or not a knee examination was performed or x-ray was taken but the results of these examinations were not captured.

Patients were excluded if they were aged younger than 30 years or older than 80 years, were unfamiliar with smartphones, or had baseline pain greater than 9 on the 11-point numeric pain rating scale (NPRS; pain ratings could range from 0, no pain, to 5, moderate pain, to 10, worst possible pain) in the target-for-treatment knee while walking on a flat surface. Patients with bilateral disease were excluded with the exception of patients who were treated in only one knee and had contralateral knee pain less than 4 on NPRS while walking on a flat surface. Patients whose baseline daily step average was less than 500 or more than 8000 as assessed during the screening and run-in phases were not eligible. Also excluded were patients who had a body mass index (BMI) greater than 35 or life expectancy less than 12 months, were currently using a wearable activity monitor or analogous device, had planned surgery on any lower extremity joint or any significant medical condition that would interfere with study participation, were chronic narcotic users, or were pregnant or breastfeeding or likely to become pregnant. The protocol did not specifically exclude patients based on use of previous intra-articular injections (cortisone or hyaluronic acid).

Study Design

All eligible patients received hylan G-F 20. Participants wore a commercially available activity monitor (Jawbone UP 24) on their wrists and were instructed to remove the monitor only during the weekly charging times and in situations where the device would be submerged in water. Patients were randomized 1:1 (stratified by site) to Group A or Group B. The randomization scheme was generated by the study sponsor and stratified by site. Sealed envelopes, numbered in an ascending order for use, were provided to each site. The envelopes were opened according to ascending sequence to ensure proper randomization.

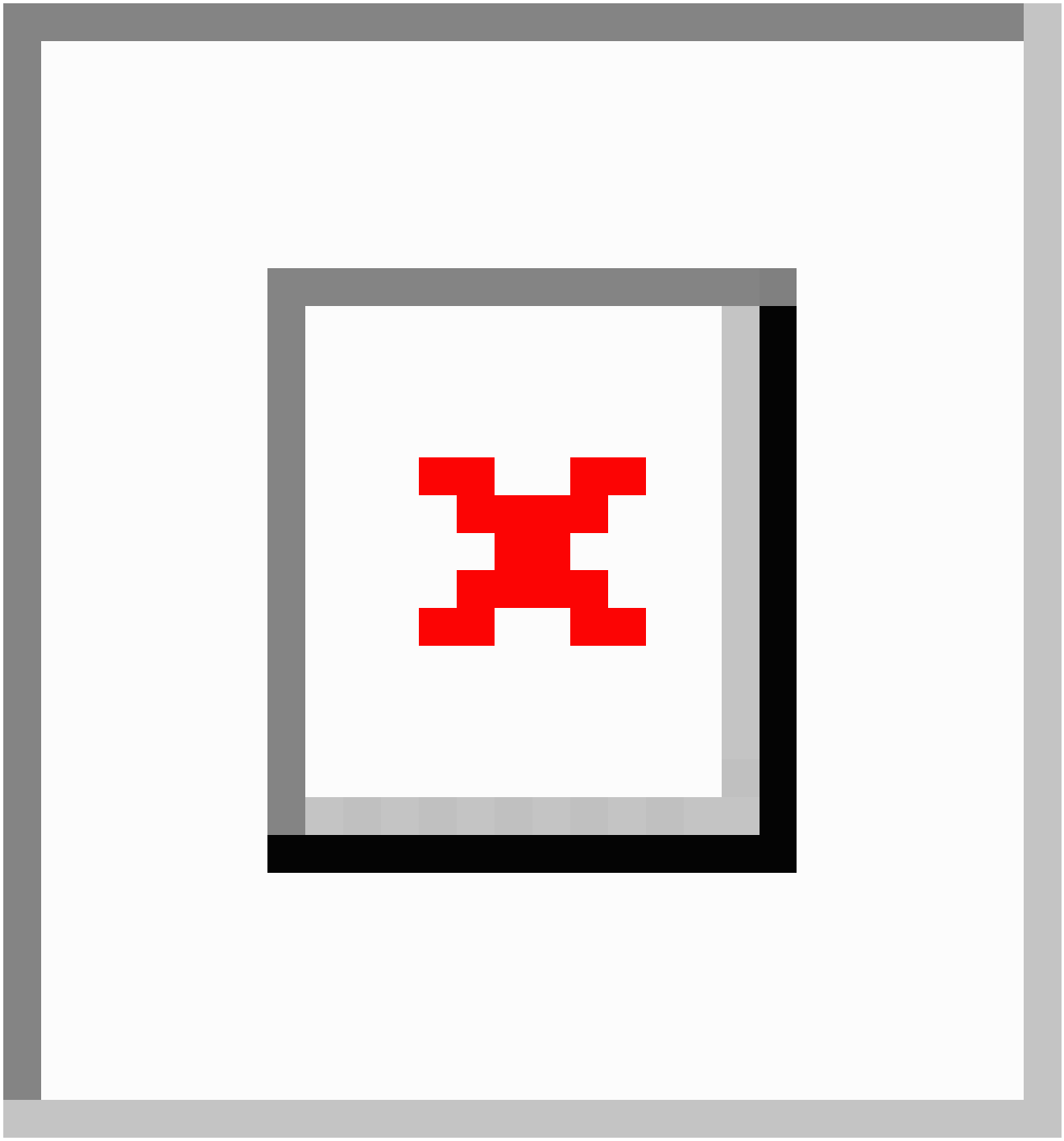
Group A patients were provided with regular follow-up as per standard-of-care (information on the benefits of walking in a brochure available from the Arthritis Foundation) plus an unblinded wearable activity monitor and a mobile app (OA GO) (Figure 1). The OA GO app (downloaded to a trial-sponsored iPhone 5 or newer) provided motivational messages and requested that the patient enter pain and mood data on a once-daily basis. Trial coordinators demonstrated app use, provided charging instructions for the Jawbone UP 24, and set the daily step goal based on patient's baseline steps per day during screening. The OA GO app combined continuously retrieved data from the wearable activity monitor with data entered by the patient to display the patient's daily step count, calories burned, and sleep. Daily and monthly cumulative activity trends were available for the patient to review. Group B was provided with regular follow-up plus a blinded wearable activity monitor, with instructions to wear the monitor at all times except during weekly charging and water activities. The patients in Group B received standard-of-care instructions and education but did not have access to activity recorded by the wearable activity monitor. In this group, data from the activity monitor were downloaded by the study team at last visit. The study consisted of 5 visits: screening and baseline (days −7 and 1) with follow-up visits at days 7 and 30, and day 90, last visit. The main study duration was 90 days. A poststudy adherence check in Group A patients occurred at day 180, when data were downloaded from the app (no visit).

Assessments

The primary endpoint was mean change from baseline to day 90 in mobility as measured by steps per day. Baseline steps per day was the average over at least 3 days, and last assessment was the average steps per day over the 7 days immediately preceding the last visit. Change was calculated as the difference between the 2 averages. Use of average steps per day avoided bias due to any one particularly good or bad day for the patient. The secondary endpoints were mean percentage change from baseline in steps per day at each assessment visit (average of a 7-day period), and at day 90, mean percentage change from baseline in the 6-minute walk test (distance and pain assessed by the NPRS), patient and physician satisfaction with treatment, percentage change in Patient Activation Measure (PAM)-13 questionnaire score [30], percentage change in sleep captured by the wearable activity monitor (light, sound, and duration of sleep), and Visual Analog Mood Scale (VAMS) assessment. Treatment-emergent adverse events (TEAEs) were also assessed.

For Group A, the satisfaction of patients and physicians was captured using survey questionnaires. The patient survey consisted of 7 questions: the first 6 were answered on a 7-point Likert scale related to the extent to which the patients experienced changes (−3 to +3, where 0=no change), and the seventh was a 4-point scale asking about their likelihood of using the device in the future (not at all likely to very likely). Valid questionnaires were those in which at least 4 of the first 6 questions were answered. A physician survey was similarly constructed with a final question, using the 4-point scale, which asked whether they would recommend the device in the future (not likely to very likely). Physicians answered the survey at day 90, before viewing the data, and must have answered at least 3 of the first 4 questions. All responses to survey questions were based on patient and physician observations and opinions. The survey questionnaires were developed for this study and were not externally validated.

Figure 1. OA GO mobile app.



The PAM-13 [30] was used to assess patients' knowledge, skills, and confidence for self-management at randomization and last visit, with responses given on a Guttman-like scale that ranged from 1 to 4 (strongly disagree to strongly agree). Unanswered items were scored as missing. Raw scores ranged from 13 to 52, with higher scores indicating more activation; raw scores were converted to activation scores, which ranged from 0 to 100, with a score of 100 corresponding to the highest degree of activation. The PAM-13 is a reliable and valid measure of patient activation with higher activation scores associated with increased self-management behaviors and increased self-efficacy with respect to patients taking charge of their overall health and involvement in their care. As a result, the concept of activation can be useful for evaluating interventions and for tailoring care plans for individual patients in a clinical setting [30,31].

The VAMS was used to assess mood as captured in 8 domains with raw scores transformed into T-scores with a mean of 50 and standard deviation of 10. The VAMS has been validated in both normal and neurologically impaired individuals as a brief measure of internal mood state [32,33].

Statistical Analysis

The Consolidated Standards of Reporting Trials guidelines were followed [34]. A sample size of 200 randomized (172 evaluable) patients (100 randomized [86 evaluable] per group) was estimated to provide 80% power to detect an average increase in the change from baseline steps per day of 25% for Group A

compared with Group B, assuming a 2-sided significance level of 5%. The variability was expected to be similar in the 2 groups, and an estimate of 58% for the coefficient of variation was used in this calculation [35]. The sample size of 200 randomized patients assumed a drop-out rate of 15%.

All efficacy endpoints were analyzed in the modified intent-to-treat population, which included all randomized patients who had both a baseline value and an on-treatment value for the primary endpoint on or after day 30. All efficacy endpoint data comparing change from baseline between groups were obtained from an analysis of covariance model with baseline mean steps per day as the covariate and treatment assignment and pooled site as class variables. They are presented as least squares (LS) means with 95% confidence intervals (CIs). Statistical analyses were based on initial verification of parametric model assumptions, and a rank analysis of covariance was then considered if conditions for a parametric model were not met. To allay concerns by the readers, nonparametric analyses are presented here. Patient and physician satisfaction surveys at last visit were summarized.

Safety endpoints were analyzed for all patients provided with a wearable activity monitor. Adverse events (AEs) were presented as the number and percentage of patients experiencing an AE. Multiple occurrences of the same event in the same patient were counted only once. The denominator for computation of percentages is the safety population within each randomized group.

Results

Patient Disposition, Baseline Characteristics, and Adherence

All patients had a knee exam at baseline, and 119 of 211 (56.4%) had an x-ray of the knee. A total of 211 knee OA patients treated with hylan G-F 20 were randomized (Group A=107; Group B=104), and nearly all patients completed the 90-day observation period (Group A, 104/107 [97.2%]; Group B, 103/104 [99.0%]) (Figure 2). Almost all Group A patients (101/107) decided to continue using the OA GO app and entered the 90 to 180 days adherence period, and 81/101 (80.2%) of these patients completed the 180 days. Baseline characteristics were similar between the 2 groups (Table 1) with a mean patient age of 62.6 (SD 9.4) years. For the total population at baseline, mean number of steps per day was 4276 (SD 1807) and pain NPRS score on the 6-minute walk test was 4.8 (SD 2.2), with 127/211 (60.2%) patients reporting pain of 5 or greater. Baseline PAM-13 score for the total population was 71.2 (SD 13.2), indicating that patients were highly activated. Overall, 91.0% of patients were compliant with the activity monitor (defined as using the wearable activity monitor 80% or more of the time). Group A had 103/107 (96.3%) patients and Group B had 80/104 (76.9%) patients who were compliant. Of the patients in Group A who entered the 90 to 180 days adherence period, 36/101 (35.6%) were 80% or more compliant with use of the OA GO app.

Figure 2. Patient disposition (a: 90-day observation period, b: 91 to 180 days).

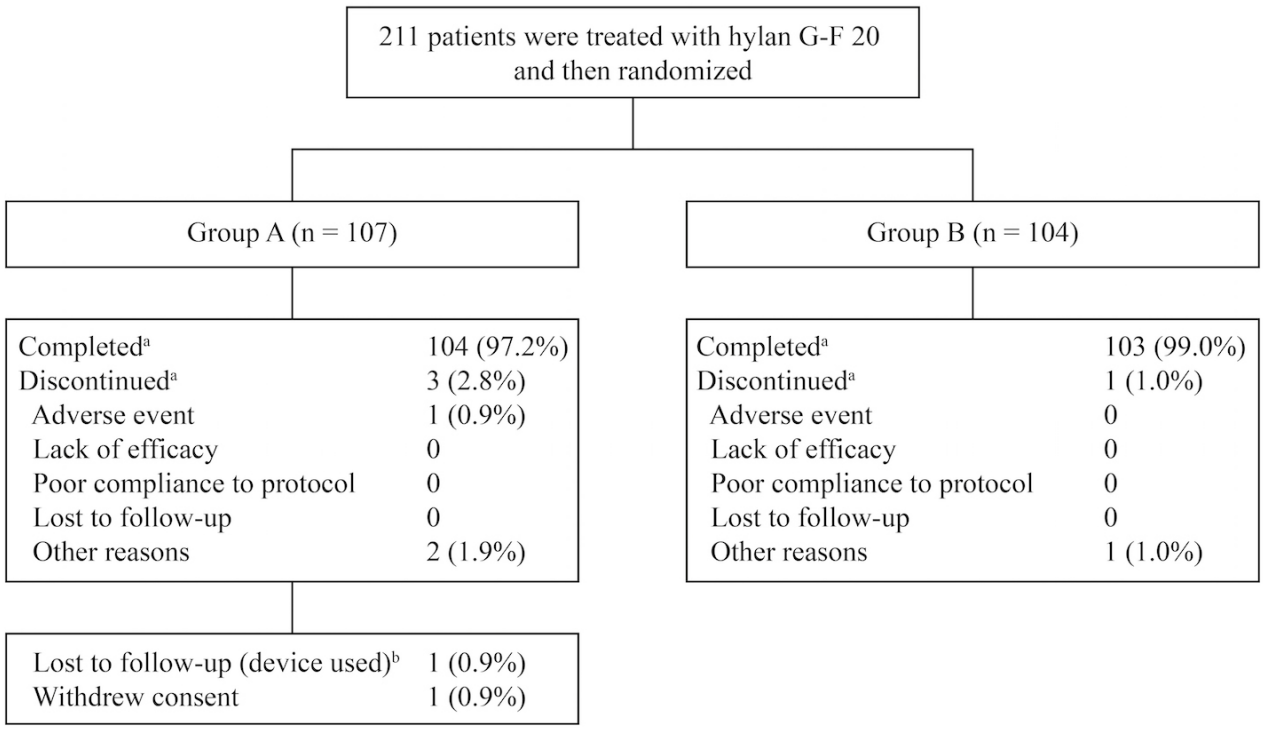


Table 1. Patient demographics and baseline characteristics.

	Group A (n=107)	Group B (n=104)	Total (N=211)
Characteristic			
Age (years), mean (SD)	61.6 (9.5)	63.6 (9.3)	62.6 (9.4)
Male, n (%)	48 (44.9)	57 (54.8)	105 (49.8)
Caucasian, n (%)	87 (81.3)	98 (94.2)	185 (87.7)
BMI[a], kg/m^2, mean (SD)	29.4 (3.9)	29.3 (3.4)	29.3 (3.7)
Steps per day, mean (SD)	4279.7 (1787.3)	4271.5 (1837.0)	4275.7 (1807.2)
6-minute Walk Test			
Pain NPRS[b], mean (SD)	4.6 (2.3)	5.1 (2.0)	4.8 (2.2)
Distance, meters, mean (SD)	402.8 (120.5)	395.6 (104.2)	399.3 (112.6)
PAM-13[c]			
Activation score, mean (SD)	71.7 (13.5)	70.6 (13.0)	71.2 (13.2)

[a]BMI: body mass index.

[b]NPRS: numeric pain rating scale.

[c]PAM-13: Patient Activation Measure-13.

Mobility and Pain

A significant increase from baseline to last assessment in mean number of steps per day was observed for all subjects (916 [SE 14.4]; *P*<.001). Improvement in mobility at day 90 was significantly greater in Group A than in Group B. LS mean change in number of steps per day was 1199 vs 467, a mean difference of 732 steps (95% CI 127-1337; *P*=.03) (Figure 3).

An increase in mean steps per day was observed for both groups at each visit. The percentage increase in steps per day was significantly greater for all visits in Group A versus Group B. LS mean percentage change in steps per day for Group A versus Group B at day 7 was 16.8% versus 3.1% (mean difference: 13.7%, 95% CI 0.7-26.7; *P*=.03), at day 30 was 40.1% versus 9.0% (mean difference: 31.1%, 95% CI 16.2-45.9; *P*<.001), and at day 90 was 35.8% versus 11.5% (mean difference: 24.3%, 95% CI 6.9-41.7; *P*=.02) (Figure 3).

In the 6-minute walk test, Group A experienced a significantly greater improvement from baseline to day 90 in LS mean percentage change in pain versus Group B, −55.3% versus −33.8% (mean difference: −21.5%, 95% CI −37.8 to −5.2; *P*=.007) (Figure 4). There was also a numerical, but not a statistically significant, improvement in LS mean percentage change in distance in the 6-minute walk test for Group A versus Group B of 18.2% versus 6.3% (mean difference: 11.9%, 95% CI −1.4 to 25.1; *P*=.96) (Figure 4).

Figure 3. Steps per day. Mean change from baseline at day 90 in number and mean percentage change from baseline. Data are presented as least squares means and standard error. P values obtained from rank analysis of covariance.

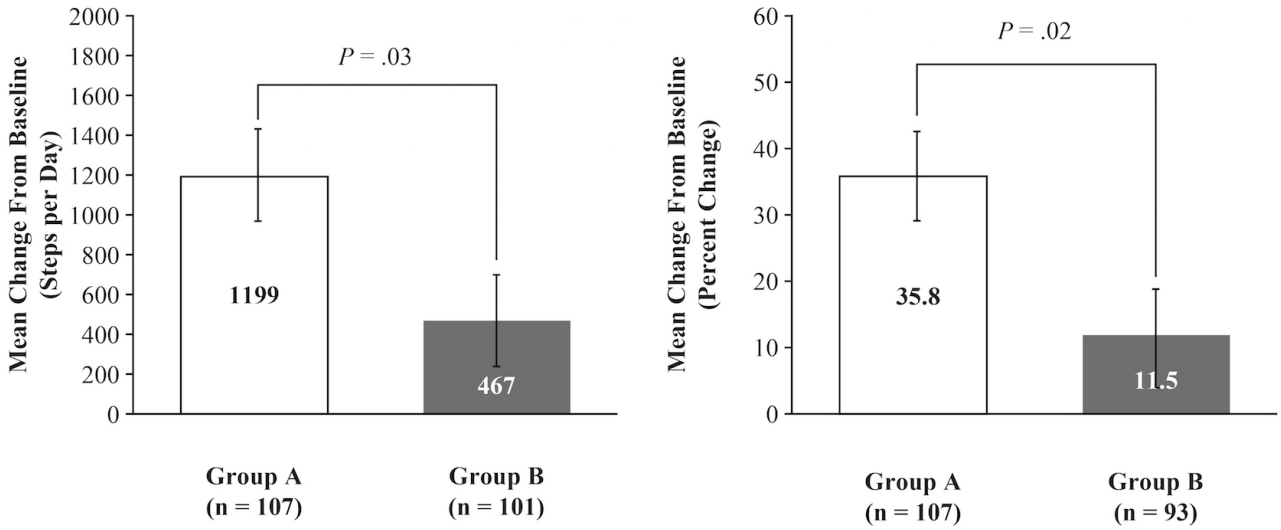
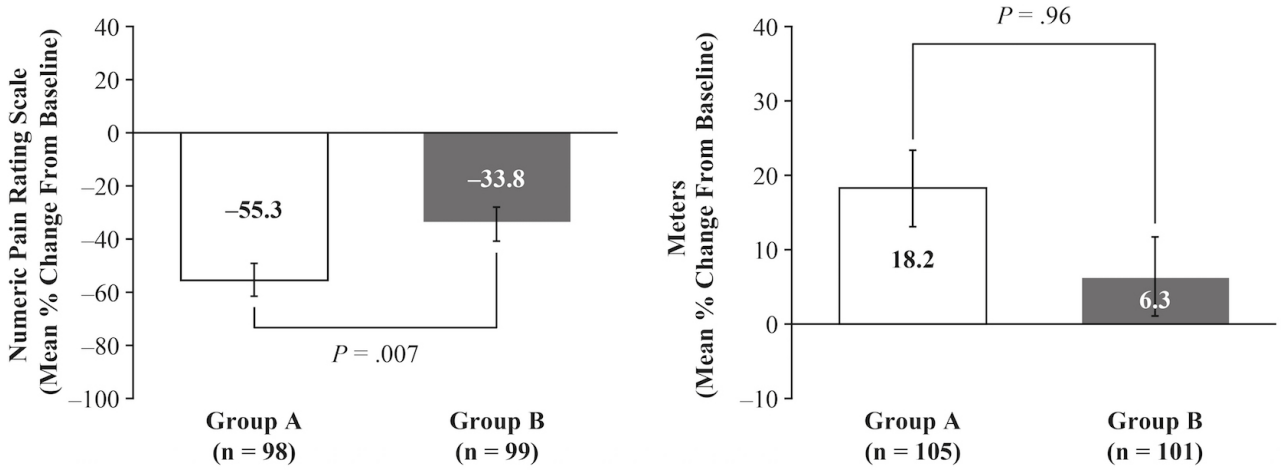


Figure 4. Six-minute walk test. Mean percentage change from baseline to day 90 in pain during the test and distance walked. Data are presented as least squares means and standard error. P values obtained from rank analysis of covariance.



Quality of Life

A greater number of Group A patients (68/104, 65.4%) reported they would be likely or very likely to use the devices compared with patients (36/104, 34.6%) who reported that they would be somewhat likely or not at all likely to do so (Figure 5). A total of 76 physicians answered surveys for 104 patients; 70 of 104 (67.3%) physician surveys reported physicians to be likely or very likely to recommend use of the devices versus 34 (32.7%) surveys reporting physicians to be only somewhat likely or not at all likely to recommend their use (Figure 5).

PAM-13 scores improved from baseline to day 90 in both groups. The LS mean change from baseline was 5.0% in Group A versus 6.9% in Group B (Figure 6; mean difference –1.9%, 95% CI –6.8% to 3.1%; *P*=.99).

There were no significant changes in sleep from baseline to day 90 for either group and no significant differences between groups. Changes in VAMS scores from baseline to day 90 were also not significantly different between groups.

Safety and Tolerability

The occurrence of TEAEs was similar in the 2 groups (Table 2) with no new safety signals noted. Arthralgia (Group A, 8/107 [7.5%]; Group B, 12/104 [11.5%]) and upper respiratory tract infection (Group A, 7/107 [6.5%]; Group B, 2/104 [1.9%]) were the 2 most commonly reported TEAEs among patients. No major AEs or treatment-emergent serious AEs related to the device or protocol occurred.

Figure 5. Satisfaction survey results in patients and physicians. Only patients in Group A (n=104) and their associated physicians participated in the satisfaction survey, which was completed at day 90.

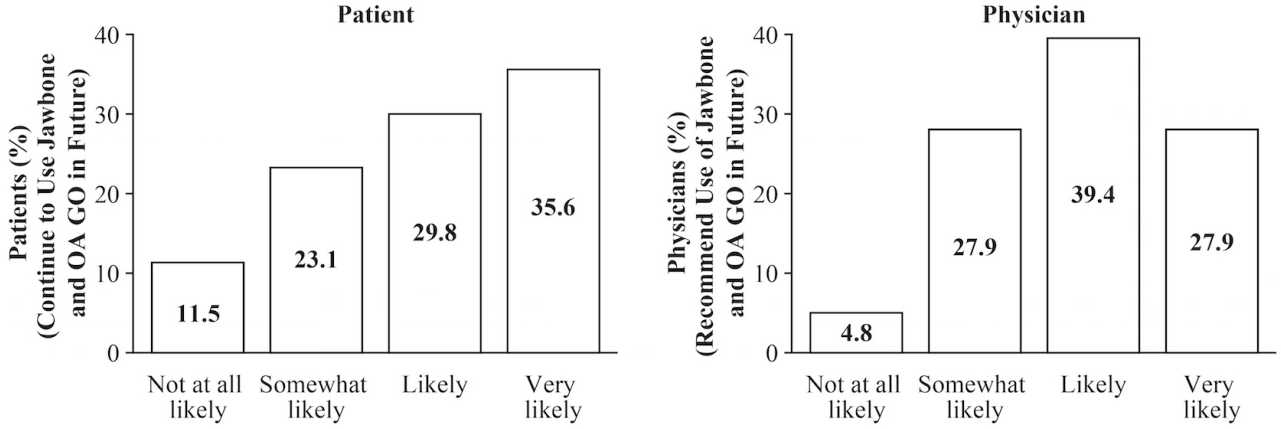


Figure 6. Mean change from baseline in Patient Activation Measure-13. Data are presented as least squares means and standard error.

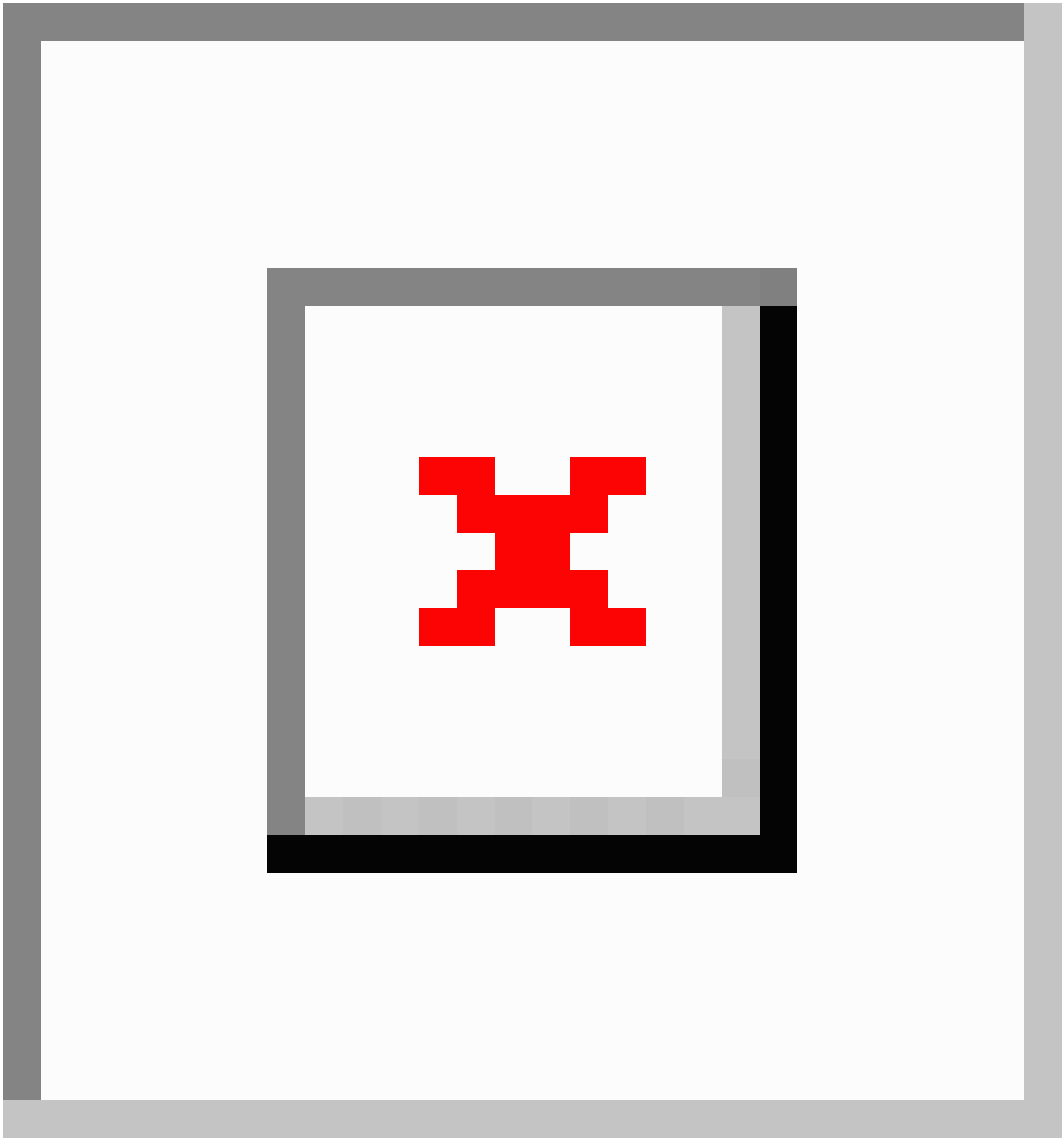


Table 2. Treatment-emergent adverse events.

	Group A (n=107)	Group B (n=104)
Any TEAE[a,b]	42 (39.3)	38 (36.5)
Any TEAE occurring in ≥2% of patients		
Arthralgia	8 (7.5)	12 (11.5)
Upper respiratory tract infection	7 (6.5)	2 (1.9)
Any serious TEAEs[c]	5 (4.7)	1 (1.0)
Any TEAE leading to death	0	0
Any TEAE leading to discontinuation of Jawbone[d]	1 (0.9)	0
Any TEAE leading to discontinuation of OA GO[d,e]	1 (0.9)	—

[a]TEAE: treatment-emergent adverse events.

[b]All TEAEs were mild to moderate in severity.

[c]Serious AEs included upper respiratory tract infection, transient ischemic attack, large intestine perforation, arthralgia and worsening OA, each in one patient in Group A and cholecystitis in one patient in Group B.

[d]Same patient discontinued both devices; the TEAE leading to discontinuation was worsening OA.

[e]OA: osteoarthritis.

Discussion

Principal Findings

Walking has been shown to have a beneficial effect on symptoms and decreases risk of functional limitations in patients with knee OA [19,20]; however, 66% of patients with arthritis reported walking for fewer than 90 minutes per week [36]. This study showed that, in patients treated with hylan G-F 20, those also using the OA GO app significantly improved their mobility with an increase in over twice as many steps per day and over 3 times the percentage change in steps per day compared with patients not using this motivating app. The increased mobility experienced by the patients using OA GO was also accompanied by a significantly greater reduction in reported pain compared with that for those not using the app. In addition, most patients and physicians expressed satisfaction with the use of the app and wearable activity monitor, suggesting that they would be amenable to the adoption of the technology in their clinical practice and daily lives. The above findings as a whole are clinically and socially important, given the recognition that increased mobility is closely associated with improved quality of life [8]. This study is the first to demonstrate that the use of a motivating mobile app can increase mobility and improve pain outcomes in patients with knee OA.

Implication of Study Findings

Use of the OA GO app provides direct feedback on mobility function and pain (whereby steps and mobility become surrogates for pain and function, assuming there are no other reasons for decreased activity), which should give a more valid and reliable report of pain over time as it does not rely on patient recall of pain experienced. In contrast, visual analog scales [37] and the Western Ontario and McMaster Universities Osteoarthritis Index (used to assess pain, stiffness, and physical function over a specific time frame [ie, the last 48 hours] in clinical trials [38]) are both subjective measures that rely on patients thinking back to how they perceived they were feeling. Furthermore, symptoms of OA are often variable depending on a number of factors, including patients' activity level [7] and changes in the weather [39], suggesting that pain should be assessed in the context of ongoing physical activity levels. This technology, therefore, is an objective measure of functional improvement over time resulting from clinical interventions provided to patients, making it not only a motivating device for patients but also an important research tool.

In this study, the results of the distance portion of the 6-minute walk test were not statistically different between groups. However, this assessment represents only a very brief snapshot in time; distance monitored continuously may be a more reliable measure of improvement in patient mobility. Despite these results not achieving statistical significance, those using the OA GO app ended the study with an average of more than 5500 steps per day, which approaches 6000 steps per day, a level that has been shown to cut the risk of developing functional limitation by half within 24 months [19].

PAM-13 scores were also not statistically different between groups in this study. The observed lack of statistical difference between groups could be due to the high mean baseline activation state of patients in this study (71.2 [SD 13.2]), which may have resulted from the selection criteria (ie, selecting relatively high functioning patients with knee OA) and may have been higher than for those excluded from the study. Indeed, the baseline activation score in this study was higher than scores for the general population in the US (61.9 [SD 21]) and populations with other chronic conditions (a US population with diabetes and with/without other comorbidities, 57.1 [40]; a Korean OA population, 56.0 [41]; and a Norwegian community mental health center population, 51.9 [42]). Evidence suggests that patients who start at the lowest activation levels tend to increase the most, suggesting that patients in this study may have been approaching a ceiling at which their ability to further

increase activation levels may have been reduced, compared with patients whose baseline activation levels were lower [43].

Limitations

This study does have some limitations. The patients studied may not be fully representative of the general OA population with respect to gender distribution (women are at higher risk for knee OA than men [44]), and activity levels used as inclusion and exclusion criteria were such that patients with low levels (less than 500 steps per day) and high levels (more than 8000 steps per day) of baseline activity were not eligible for enrollment. Patients with a BMI above 35 were also excluded from the study. In addition, the Kellgren-Lawrence grade of patients' knee OA was not collected, so data could not be stratified and analyzed in relation to radiographic disease severity, which might have been helpful in clarifying some of the study results such as the PAM-13 scores. The patient and physician satisfaction survey collected important feedback but is not an externally validated instrument. Moreover, the study was short term, having been conducted over 90 days. Finally, the study did not include an arm in which patients used the motivating app and did not receive an injection, nor did it include an arm in which patients used the motivating app and received a saline injection (often designated as placebo).

Practical Considerations in the Use of Mobile Health Apps

The practical aspects of incorporating a mobile health app into treatment paradigms to improve patient mobility should be further investigated. This strategy requires access to a smartphone, which may be a financial barrier for some patients [45]. Patients must also understand how the smartphone app works and have the ability and confidence to use it effectively [45]. Finally, with the ever-increasing complexity of mobile apps, continual improvements in smartphones, apps, and wearable activity monitors are needed in order to limit or prevent technical issues (and potential data loss) that may occur with routine operations such as software updates [46,47]. Continual improvement helps ensure that data are captured accurately while avoiding false positives and negatives [48,49].

Overall, the results of this study provide evidence that in patients suffering from knee OA who received hylan G-F 20, additional improvement was achieved in this study, based on mobility parameters of steps per day and pain reduction, with use of a mobile app and wearable activity monitor.

Conclusions

Reduced mobility in patients with knee OA can be a significant issue negatively affecting quality of life and experience of pain. Increasing patients' motivation to walk and thereby increasing their mobility may reduce these negative effects. Use of a novel smartphone app in conjunction with a wearable activity monitor provided additional improvement on mobility parameters such as steps per day and pain with walking in the 6-minute walk test in patients with knee OA who were treated with hylan G-F 20 (see Multimedia Appendix 1). These results also highlight the amenability of patients and physicians to using mobile health technology in the treatment of OA and suggest further investigation is warranted.

Acknowledgments

Sanofi Biosurgery LLC funded the study. Assistance with the writing and development of the manuscript was provided by Nicholas C Stilwell, PhD, and Evidence Scientific Solutions and was funded by Sanofi Biosurgery LLC. The authors would like to acknowledge Bruno Leroy, MD, of Sanofi who contributed to the trial concept and study design; Lynn Mcroy, MD, who contributed to the trial concept and design and OA GO app development; and Scott D Kelly, who was an employee of Sanofi when the study was conducted and contributed to data analysis. The authors would also like to thank the following investigators: Ladan Bakhtari, MD, Doctors of Internal Medicine Ltd/Clinical Research Advantage, Plano, TX; John Ervin, MD, The Center for Pharmaceutical Research, Kansas City, MO; Richard Glover II, MD, Heartland Research Associates, LLC, Newton, KS; B Matthew Hicks, MD, Fort Wayne Orthopedics, Fort Wayne IN; Jeffry Jacqmein, MD, Jacksonville Center for Clinical Research, Jacksonville, FL; Ted Jagielo, MD, Medical and Procedural Specialists of Illinois/Clinical Research Advantage, Inc, Chicago, IL; David Kirby, MD, Clinical Research Advantage, Birmingham, AL; Tracy Klein, MD, Heartland Research Associates, LLC, Wichita, KS; Louis Levy, MD, TriWest Research Associates, LLC, El Cajon, CA; Barry Lubin, MD, National Clinical Research Center, Norfolk, VA; Erich Schramm, MD, St. John's Center for Clinical Research, Ponte Vedra, FL; Tammi Shlotzhauer, MD, Rochester Clinical Research, Rochester, NY; Michael Skyhar, MD, CORE Orthopaedic Medical Center, Encinitas, CA; Cynthia Strout, MD, Coastal Caroline Research Center, Mt Pleasant, SC; and L Tyler Wadsworth, MD, Sundance Clinical Research, Norfolk, VA.

Conflicts of Interest

Nebojsa Skrepnik has been a consultant for Sanofi and Ortofix. Andrew Spitzer has been a consultant for DePuy, Flexion Therapeutics, and Sanofi-Aventis; has received research support from DePuy; and has been a speaker for Sanofi-Aventis. Roy Altman has been a consultant for Cytori, DuPuy, Ferring, Flexion, Iroko, McNeil, Novartis, GlaxoSmithKline, Olatec, Pfizer, Q-Med, Rotta (MEDA), Strategic Science & Technologies, and Teva and has been a speaker for Iroko. John Hoekstra was a principal investigator in this study while employed at National Clinical Research Inc. John Stewart is an employee of Sanofi. Richard Toselli was an employee of Sanofi at the time the work was conducted and is currently an employee of Cochlear Ltd, Sydney, Australia.

Multimedia Appendix 1

Patient experience of smartphone app use during the study.

[MOV File, 73MB - mhealth_v5i5e64_app1.mov]

Multimedia Appendix 2

CONSORT-EHEALTH-v1-6.

[PDF File (Adobe PDF File), 12MB - mhealth_v5i5e64_app2.pdf]

References

1. Murphy L, Helmick CG. The impact of osteoarthritis in the United States: a population-health perspective. Am J Nurs 2012 Mar;112(3 Suppl 1):S13-S19. [doi: 10.1097/01.NAJ.0000412646.80054.21] [Medline: 22373741]
2. Centers for Disease Control and Prevention. Prevalence of doctor-diagnosed arthritis and arthritis-attributable activity limitation—United States, 2010-2012. MMWR Morb Mortal Wkly Rep 2013 Nov 08;62(44):869-873 [FREE Full text] [Medline: 24196662]
3. Johnson VL, Hunter DJ. The epidemiology of osteoarthritis. Best Pract Res Clin Rheumatol 2014 Feb;28(1):5-15. [doi: 10.1016/j.berh.2014.01.004] [Medline: 24792942]
4. Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, National Arthritis Data Workgroup. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum 2008 Jan;58(1):26-35 [FREE Full text] [doi: 10.1002/art.23176] [Medline: 18163497]
5. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Clin Geriatr Med 2010 Aug;26(3):355-369 [FREE Full text] [doi: 10.1016/j.cger.2010.03.001] [Medline: 20699159]
6. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. Lancet 2011 Jun 18;377(9783):2115-2126. [doi: 10.1016/S0140-6736(11)60243-2] [Medline: 21684382]
7. Hawker GA, Stewart L, French MR, Cibere J, Jordan JM, March L, et al. Understanding the pain experience in hip and knee osteoarthritis—an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008 Apr;16(4):415-422 [FREE Full text] [doi: 10.1016/j.joca.2007.12.017] [Medline: 18296075]
8. Fagerström C, Borglin G. Mobility, functional ability and health-related quality of life among people of 60 years or older. Aging Clin Exp Res 2010;22(5-6):387-394. [Medline: 21422794]
9. Hubertsson J, Petersson IF, Thorstensson CA, Englund M. Risk of sick leave and disability pension in working-age women and men with knee osteoarthritis. Ann Rheum Dis 2013 Mar;72(3):401-405. [doi: 10.1136/annrheumdis-2012-201472] [Medline: 22679305]
10. Kotlarz H, Gunnarsson CL, Fang H, Rizzo JA. Osteoarthritis and absenteeism costs: evidence from US National Survey Data. J Occup Environ Med 2010 Mar;52(3):263-268. [doi: 10.1097/JOM.0b013e3181cf00aa] [Medline: 20190656]
11. Reginster J, Reiter-Niesert S, Bruyère O, Berenbaum F, Brandi M, Branco J, et al. Recommendations for an update of the 2010 European regulatory guideline on clinical investigation of medicinal products used in the treatment of osteoarthritis and reflections about related clinically relevant outcomes: expert consensus statement. Osteoarthritis Cartilage 2015 Dec;23(12):2086-2093 [FREE Full text] [doi: 10.1016/j.joca.2015.07.001] [Medline: 26187570]
12. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, American College of Rheumatology. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2012 Apr;64(4):465-474. [Medline: 22563589]
13. Jordan KM, Arden NK, Doherty M, Bannwarth B, Bijlsma JW, Dieppe P, Standing Committee for International Clinical Studies Including Therapeutic Trials ESCISIT. EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). Ann Rheum Dis 2003 Dec;62(12):1145-1155 [FREE Full text] [Medline: 14644851]
14. Richmond J, Hunter D, Irrgang J, Jones MH, Levy B, Marx R, American Academy of Orthopaedic Surgeons. Treatment of osteoarthritis of the knee (nonarthroplasty). J Am Acad Orthop Surg 2009 Sep;17(9):591-600 [FREE Full text] [Medline: 19726743]
15. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008 Feb;16(2):137-162 [FREE Full text] [doi: 10.1016/j.joca.2007.12.013] [Medline: 18279766]
16. Marks R. Knee osteoarthritis and exercise adherence: a review. Curr Aging Sci 2012 Feb;5(1):72-83. [Medline: 21762086]
17. Pisters MF, Veenhof C, Schellevis FG, Twisk JW, Dekker J, De Bakker DH. Exercise adherence improving long-term patient outcome in patients with osteoarthritis of the hip and/or knee. Arthritis Care Res (Hoboken) 2010 Aug;62(8):1087-1094 [FREE Full text] [doi: 10.1002/acr.20182] [Medline: 20235201]
18. Dobson F, Hinman RS, Hall M, Terwee CB, Roos EM, Bennell KL. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: a systematic review. Osteoarthritis Cartilage 2012 Dec;20(12):1548-1562 [FREE Full text] [doi: 10.1016/j.joca.2012.08.015] [Medline: 22944525]

19. White DK, Tudor-Locke C, Zhang Y, Fielding R, LaValley M, Felson DT, et al. Daily walking and the risk of incident functional limitation in knee osteoarthritis: an observational study. Arthritis Care Res (Hoboken) 2014 Sep;66(9):1328-1336 [FREE Full text] [doi: 10.1002/acr.22362] [Medline: 24923633]
20. Evcik D, Sonel B. Effectiveness of a home-based exercise therapy and walking program on osteoarthritis of the knee. Rheumatol Int 2002 Jul;22(3):103-106. [doi: 10.1007/s00296-002-0198-7] [Medline: 12111084]
21. Flores MG, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
22. Granado-Font E, Flores-Mateo G, Sorlí-Aguilar M, Montaña-Carreras X, Ferre-Grau C, Barrera-Uriarte M, et al. Effectiveness of a smartphone application and wearable device for weight loss in overweight or obese primary care patients: protocol for a randomised controlled trial. BMC Public Health 2015 Jun 04;15:531 [FREE Full text] [doi: 10.1186/s12889-015-1845-8] [Medline: 26041131]
23. Guy R, Hocking J, Wand H, Stott S, Ali H, Kaldor J. How effective are short message service reminders at increasing clinic attendance? A meta-analysis and systematic review. Health Serv Res 2012 Apr;47(2):614-632 [FREE Full text] [doi: 10.1111/j.1475-6773.2011.01342.x] [Medline: 22091980]
24. Weinberger M, Tierney WM, Booher P, Katz BP. Can the provision of information to patients with osteoarthritis improve functional status? A randomized, controlled trial. Arthritis Rheum 1989 Dec;32(12):1577-1583. [Medline: 2688659]
25. Nguyen HQ, Gill DP, Wolpin S, Steele BG, Benditt JO. Pilot study of a cell phone-based exercise persistence intervention post-rehabilitation for COPD. Int J Chron Obstruct Pulmon Dis 2009;4:301-313 [FREE Full text] [Medline: 19750190]
26. Antypas K, Wangberg SC. An Internet- and mobile-based tailored intervention to enhance maintenance of physical activity after cardiac rehabilitation: short-term results of a randomized controlled trial. J Med Internet Res 2014;16(3):e77 [FREE Full text] [doi: 10.2196/jmir.3132] [Medline: 24618349]
27. Hong YA, Goldberg D, Ory MG, Towne SD, Forjuoh SN, Kellstedt D, et al. Efficacy of a mobile-enabled Web app (iCanFit) in promoting physical activity among older cancer survivors: a pilot study. JMIR Cancer 2015 Jun 26;1(1):e7 [FREE Full text] [doi: 10.2196/cancer.4389] [Medline: 28410158]
28. Poirier J, Bennett WL, Jerome GJ, Shah NG, Lazo M, Yeh H, et al. Effectiveness of an activity tracker- and Internet-based adaptive walking program for adults: a randomized controlled trial. J Med Internet Res 2016;18(2):e34 [FREE Full text] [doi: 10.2196/jmir.5295] [Medline: 26860434]
29. Safran NJ, Madar Z, Shahar DR. The impact of a Web-based app (eBalance) in promoting healthy lifestyles: randomized controlled trial. J Med Internet Res 2015;17(3):e56 [FREE Full text] [doi: 10.2196/jmir.3682] [Medline: 25732936]
30. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health Serv Res 2005 Dec;40(6 Pt 1):1918-1930 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00438.x] [Medline: 16336556]
31. 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]
32. Stern RA, Arruda JE, Hooper CR, Wolfner GD, Morey CE. Visual analogue mood scales to measure internal mood state in neurologically impaired patients: Description and initial validity evidence. Aphasiology 1997 Jan;11(1):59-71. [doi: 10.1080/02687039708248455]
33. Arruda JE, Stern RA, Somerville JA. Measurement of mood states in stroke patients: validation of the visual analog mood scales. Arch Phys Med Rehabil 1999 Jun;80(6):676-680. [Medline: 10378494]
34. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010 Jun 1;152(11):726-732. [doi: 10.7326/0003-4819-152-11-201006010-00232] [Medline: 20335313]
35. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
36. Centers for Disease Control and Prevention (CDC). State-specific prevalence of walking among adults with arthritis—United States, 2011. MMWR Morb Mortal Wkly Rep 2013 May 03;62(17):331-334 [FREE Full text] [Medline: 23636026]
37. Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. Res Nurs Health 1990 Aug;13(4):227-236. [Medline: 2197679]
38. Bellamy N. WOMAC Osteoarthritis Index: WOMAC 3.1 Index. URL: http://www.womac.com/womac/index.htm [accessed 2016-07-25] [WebCite Cache ID 6jGuzl5mH]
39. McAlindon T, Formica M, Schmid CH, Fletcher J. Changes in barometric pressure and ambient temperature influence osteoarthritis pain. Am J Med 2007 May;120(5):429-434. [doi: 10.1016/j.amjmed.2006.07.036] [Medline: 17466654]
40. Remmers C, Hibbard J, Mosen DM, Wagenfield M, Hoye RE, Jones C. Is patient activation associated with future health outcomes and healthcare utilization among patients with diabetes? J Ambul Care Manage 2009;32(4):320-327. [doi: 10.1097/JAC.0b013e3181ba6e77] [Medline: 19888008]

41. Ahn YH, Kim BJ, Ham OK, Kim SH. Factors associated with patient activation for self-management among community residents with osteoarthritis in Korea. J Korean Acad Community Health Nurs 2015;26(3):303. [doi: 10.12799/jkachn.2015.26.3.303]
42. Moljord IE, Lara-Cabrera ML, Perestelo-Pérez L, Rivero-Santana A, Eriksen L, Linaker OM. Psychometric properties of the Patient Activation Measure-13 among out-patients waiting for mental health treatment: A validation study in Norway. Patient Educ Couns 2015 Nov;98(11):1410-1417 [FREE Full text] [doi: 10.1016/j.pec.2015.06.009] [Medline: 26146239]
43. 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]
44. Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. Osteoarthritis Cartilage 2005 Sep;13(9):769-781 [FREE Full text] [doi: 10.1016/j.joca.2005.04.014] [Medline: 15978850]
45. Boulos MN, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomed Eng Online 2011;10:24 [FREE Full text] [doi: 10.1186/1475-925X-10-24] [Medline: 21466669]
46. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
47. 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]
48. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA 2015 Feb 10;313(6):625-626. [doi: 10.1001/jama.2014.17841] [Medline: 25668268]
49. 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]

Abbreviations

AE: adverse events
FDA: Food and Drug Administration
LS: least square
NPRS: numeric pain rating scale
OA: osteoarthritis
PAM-13: Patient Activation Measure-13
TEAE: treatment-emergent adverse events
VAMS: visual analog mood assessment

Edited by G Eysenbach; submitted 19.12.16; peer-reviewed by G Jerome, M Iversen; comments to author 09.01.17; revised version received 27.02.17; accepted 23.03.17; published 09.05.17.

Please cite as:
Skrepnik N, Spitzer A, Altman R, Hoekstra J, Stewart J, Toselli R
Assessing the Impact of a Novel Smartphone Application Compared With Standard Follow-Up on Mobility of Patients With Knee Osteoarthritis Following Treatment With Hylan G-F 20: A Randomized Controlled Trial
JMIR Mhealth Uhealth 2017;5(5):e64
URL: http://mhealth.jmir.org/2017/5/e64/
doi:10.2196/mhealth.7179
PMID:28487266

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org