JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 8 (2020), Issue 3 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Original Papers

Reviews

Original Paper

The Arabic Version of the Mobile App Rating Scale: Development and Validation Study

Marco Bardus[1], BA, MA, PhD; Nathalie Awada[1], MPH; Lilian A Ghandour[2], MPH, PhD; Elie-Jacques Fares[3], PhD; Tarek Gherbal[4], DPT, MSc; Tasnim Al-Zanati[5], BSc; Stoyan R Stoyanov[6], MRes (Psych)

[1]Department of Health Promotion and Community Health, Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon

[2]Department of Epidemiology and Population Health, Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon

[3]Department of Nutrition and Food Sciences, Faculty of Agricultural and Food Sciences, American University of Beirut, Beirut, Lebanon

[4]University Sports, Office of Student Affairs, American University of Beirut, Beirut, Lebanon

[5]International Medical Corps, Amman, Jordan

[6]School of Design, Creative Industries Faculty, Queensland University of Technology, Brisbane, Australia

Corresponding Author:
Marco Bardus, BA, MA, PhD
Department of Health Promotion and Community Health
Faculty of Health Sciences
American University of Beirut
Van Dyck - Room 302
PO Box 11-0236 Riad El-Solh
Beirut, 1107 2020
Lebanon
Phone: 961 1350000 ext 4694
Email: marco.bardus@gmail.com

Abstract

Background: With thousands of health apps in app stores globally, it is crucial to systemically and thoroughly evaluate the quality of these apps due to their potential influence on health decisions and outcomes. The Mobile App Rating Scale (MARS) is the only currently available tool that provides a comprehensive, multidimensional evaluation of app quality, which has been used to compare medical apps from American and European app stores in various areas, available in English, Italian, Spanish, and German. However, this tool is not available in Arabic.

Objective: This study aimed to translate and adapt MARS to Arabic and validate the tool with a sample of health apps aimed at managing or preventing obesity and associated disorders.

Methods: We followed a well-established and defined "universalist" process of cross-cultural adaptation using a mixed methods approach. Early translations of the tool, accompanied by confirmation of the contents by two rounds of separate discussions, were included and culminated in a final version, which was then back-translated into English. Two trained researchers piloted the MARS in Arabic (MARS-Ar) with a sample of 10 weight management apps obtained from Google Play and the App Store. Interrater reliability was established using intraclass correlation coefficients (ICCs). After reliability was ascertained, the two researchers independently evaluated a set of additional 56 apps.

Results: MARS-Ar was highly aligned with the original English version. The ICCs for MARS-Ar (0.836, 95% CI 0.817-0.853) and MARS English (0.838, 95% CI 0.819-0.855) were good. The MARS-Ar subscales were highly correlated with the original counterparts (P<.001). The lowest correlation was observed in the area of usability (r=0.685), followed by aesthetics (r=0.827), information quality (r=0.854), engagement (r=0.894), and total app quality (r=0.897). Subjective quality was also highly correlated (r=0.820).

Conclusions: MARS-Ar is a valid instrument to assess app quality among trained Arabic-speaking users of health and fitness apps. Researchers and public health professionals in the Arab world can use the overall MARS score and its subscales to reliably evaluate the quality of weight management apps. Further research is necessary to test the MARS-Ar on apps addressing various health issues, such as attention or anxiety prevention, or sexual and reproductive health.

(JMIR Mhealth Uhealth 2020;8(3):e16956) doi:10.2196/16956

KEYWORDS

validation studies as topic; mHealth; mobile app rating scale; Arab world; eHealth; app quality; app evaluation; mobile app

Introduction

Background

Preventing noncommunicable diseases (NCDs) is a major public health priority [1], globally and in the Arab region, where heart disease, diabetes, hypertension, stroke, and other cardiovascular disorders are commonly observed in both low-income and high-income countries [2]. The prevalence of overweight ranged from 19% to 57% in the Middle East and North Africa (MENA) region, and from 6% to 53% in the Eastern Mediterranean area [3], but it reached higher levels in high-income countries of the Gulf, such as Kuwait and the United Arab Emirates [4]. Similar trends are observed for type 2 diabetes (an estimated 9% of the population), which is projected to affect 60 million Arabs in 2030 [5].

Mobile apps provide a unique opportunity to address NCDs worldwide [6,7], as these technologies are available among both high- and low-income populations [8]. In the world, there are more than 7 billion mobile subscribers [9] (3.4 billion of whom are mobile phone users) [10]. Recent systematic reviews provide some evidence of the efficacy of mobile health (mHealth) apps for promoting dietary self-regulation [11] and weight management [12-18]. In 2017, there were more than 350,000 mHealth apps available in Web-based stores [19], offering a wide variety of services for primary or secondary prevention [20]. The global health app market was worth US $25 billion in 2017 and US $37 billion in 2019, and it is projected to reach US $72 billion in 2020 [21]. In the Arab world, the mHealth market is also rapidly growing and is expected to reach US $1.3 billion by 2019 [22]. However, the market is extremely volatile and unstable; in some cases, app turnover can be 3.7 days in Google Play (for Android phones) and 13.7 days in App Store (for iOS phones) over 9 months [23]. Some research shows that many apps are downloaded less than 500 times, or never used [24]. Qualitative studies show that users stop using health apps because of hidden costs, increased data entry burden [25], and low engagement [26]. From a content point of view, apps generally lack evidence-based and theoretical support [27,28]. The instability and unpredictability of the health app market pose several challenges for both experts (ie, health professionals and researchers) and laypersons (ie, customers, end users, and patients), who need appropriate tools to decide which apps are worth using and recommending.

Evaluating app quality has become a fundamental task for researchers, as the failure to accurately and adequately evaluate health app quality might compromise end users' well-being and decrease their confidence in the technology [23]. Various frameworks and tools exist to evaluate app quality [29], but they generally lack multidimensionality and cultural flexibility, focusing on either information content, functionality, usability, accountability, impact, or popularity dimensions [29,30].

The Mobile App Rating Scale (MARS) [31] is a multidimensional comprehensive tool for assessing the quality of mHealth apps for experts. According to the scale developers, MARS includes 19 questions or items, which have been logically grouped according to *objective* dimensions of engagement (five items), functionality (four items), aesthetics (three items), and information quality (seven items). The instrument also includes four items that are deemed more *subjective* as they include questions such as the following: "Would you recommend this app to people who might benefit from it?" "How many times do you think you would use this app in the next 12 months if it was relevant to you?" "Would you pay for this app?" and "What is your overall 5-star rating of the app?"

In the development of MARS, the authors involved a multidisciplinary team of designers, health professionals, and developers [31], making the scale user friendly, dependable, and broadly applicable to different health apps. MARS has been used by trained raters to evaluate apps addressing a wide range of behaviors and health-related issues, such as drunk-driving prevention [32], speech sound disorders [33], self-care management of heart failure symptoms [34], mental health and mindfulness [35], quality of life [36], weight loss and smoking cessation [37], or weight management, including physical activity and calorie counting apps [38]. A simplified version for end users (user version of the MARS, uMARS) has also been developed [39]; it includes the same domains of the MARS tool, using simplified language and omitting items that would require rater expertise, so that it can be used without training and by laypersons or end users [31].

The MARS tool has been recently translated into Italian [40], Spanish [41], and German [42], and there are ongoing projects for translating it into nine other languages. However, there is currently no instrument for assessing the quality of health apps in Arabic. The Arab world geographically includes Africa (Algeria, Comoros, Djibouti, Egypt, Libya, Mauritania, Morocco, Somalia, Sudan, and Tunisia), Middle East, and parts of Asia (Bahrain, Iraq, Jordan, Kuwait, Lebanon, Oman, Palestine/Israel, Qatar, Saudi Arabia, Syria, the United Arab Emirates, and Yemen). Even though the original MARS tool could be used by Arabs who are also fluent in English, the majority of people living in the MENA region have "very low" English proficiency, according to the Education First English Proficiency Index [43]. With a growing mHealth market in the Arab world, along with growing public health concerns about NCD trends in the region, there is an urgent need for tools such as MARS to be available for Arabic-speaking health professionals and end users in the region.

Objectives

This study aimed to fill the gap by adapting the MARS in Arabic (MARS-Ar) and validating the instrument with a sample of popular weight management apps, available in the category "Health and Fitness" in the app stores of the Arab world.

Methods

Study Design Overview

This study followed a well-established and so called "universalist approach" [44], which is based on the assumption that an individual's response to any given question or concept depends on the individual's culture [45]. We followed a similar procedure used by researchers who developed and validated the MARS tool in Italian [40] and German [42]. This process comprises several phases, including (1) translation and cultural adaptation with back-translation, (2) review, (3) piloting, and (4) validation or psychometric evaluation. The local Institutional Review Board approved the study protocol and research procedures involving human subjects on November 1, 2018 (ref. nr: SBS-2018-0394). In the section below, we describe the process of translation and cultural adaptation, including the review and piloting phases. In the results section, we describe the results of the validation or psychometric evaluation of the MARS-Ar tool.

Phase 1: Translation and Cultural Adaptation Process

The MARS tool was first translated in Arabic by a professional English-Arabic translator, with expertise in technological topics, who was recruited from a pool of contractors of the American University of Beirut. The translated instrument was broken down into sections and parts, including titles, introductory paragraphs and instructions, and the actual MARS items, with several answer options. MARS was segmented into 59 parts; the translated parts were laid out in a table with the original English version. Each segment was associated with a unique identifier (see Figure 1) so that it would be easier to identify any editing modifications and quantitative ratings for the translation provided by experts.

Figure 1. Format of the document used in the Mobile App Rating Scale-Arabic translation process.

Section	App Classification	
Language	English (original)	Arabic
Part ID		
MAC1	Mobile Application Rating Scale (MARS)	مقياس تقييم تطبيق هاتف جوال
MAC2	App Classification	تصنيف التطبيق
MAC3	The Classification section is used to collect descriptive and technical information about the app. Please review the app description in iTunes / Google Play to access this information.	القسم المتعلق بالتصنيف يستخدم لجمع معلومات وصفية وتقنية حول التطبيق. رجاء راجع وصف التطبيق في مخزن ال iTunes/ Google Play للوصول إلى تلك المعلومات
MAC4	App Name: _____ Rating this version: _____ Rating all versions: _____ Developer: _____ Number of ratings this version: _____ Number of ratings all versions: _____ Version: _____ Last update: _____ Cost - basic version: _____ Cost - upgrade version: _____ Platform: iPhone iPad Android Brief description: _____	اسم التطبيق:..... تقييم هذه النسخة :..... تقييم كافة النسخ:..... مطور التطبيق:..... عدد مَن قيموا هذه النسخة:..... عدد مَن قيموا كافة النسخ:..... النسخة:..... التحديث الأخير:..... الكلفة- النسخة الأساسية:..... الكلفة- النسخة المحدثة:..... المنصة: iPhone iPad Android وصف مختصر:
MAC5	Focus: what the app targets (select all that apply) o Increase Happiness/Well-being o Mindfulness/Meditation/Relaxation o Reduce negative emotions o Depression o Anxiety/Stress o Anger o Behaviour Change o Alcohol /Substance Use o Goal Setting o Entertainment o Relationships o Physical health o Other	مركز الاهتمام: ما هي الأشياء التي يستهدفها التطبيق (اختر كل ما ينطبق) - زيادة السعادة/ الرفاه - اليقظة الذهنية/ التأمل/الاسترخاء - الحد من المشاعر السلبية - الاكتئاب - القلق/الضغط - الغضب - تغيير السلوك - الكحول/ الإدمان - تحديد الأهداف - الترفيه - العلاقات - الصحة الجسدية - غيره
MAC6	Theoretical background/Strategies (all that apply) o Assessment o Feedback o Information/Education o Monitoring/Tracking	خلفيات نظرية/ استراتيجيات (كل ما ينطبق) - التقويم - التغذية الراجعة - المعارف/ التعليم - الرقابة/المتابعة

Phase 2: Review

The review phase comprised two rounds of Web-based consultations among Arabic-speaking experts from various academic and governmental institutions in the MENA region, who responded to an initial call for Arabic-speaking academics (language experts, social scientists, computer scientists, and engineers), practitioners, or app developers who would be willing to evaluate and provide feedback on the Arabic translation of MARS.

Recruitment

The research team members sent email invitations to their personal social networks and to the Public Health in the Arab World mailing list, a subscription-based email list that focuses on issues related to public health in the Arab World and includes more than 1900 subscribers worldwide. The call was also shared on professional social networking sites (eg, LinkedIn and ResearchGate) and on the research team members' personal social media profiles on Facebook and Twitter. The email and the social media posts contained a link to a consent form, stored on MailChimp servers, where interested participants provided consent for participation in the study.

Between March 26, 2019, and April 17, 2019, 19 Arabic-speaking experts from various academic and governmental institutions responded to the call and agreed to participate in the translation and cultural adaptation phase of the project. Participants included 9 representatives from Lebanon (the Ministry of Public Health, the American University of

Beirut, the Lebanese American University, and a local Nongovernmental Organization), 2 representatives from Egypt (Alexandria Regional Centre for Women's Health and Development and Egypt Health Foundation), 2 representatives from Jordan (King Hussein Cancer Center and a tech company ISEET), and 1 representative each from Syria (Action Against Hunger), Morocco (Faculty of Sciences, University Ibn Tofail, Kénitra), Qatar (Hamad Bin Khalifa University-College of Science and Engineering), Saudi Arabia (Saudi Center for Disease Control and Prevention), the United Arab Emirates (Specialized rehabilitation hospital and Capital Health), and the United States (Wayne State University).

Review Consultation Procedures

The research team set up a Web-based consultation system based on email communications through MailChimp, Google Docs, and a Web-based survey hosted on the American University of Beirut servers (LimeSurvey, GmbH) [46]. Enrolled experts received an email with a Word document containing the translation and original version of the MARS tool, as shown in Figure 1. The experts were instructed to (1) download the Word document on their computer, (2) add comments and edits to the file using "track changes," (3) upload the edited document on LimeSurvey using personalized credentials, and (4) complete an evaluation form rating the translation for each part. Experts were asked to rate the appropriateness and accuracy of each segment using 5-point Likert-type scales (5=very appropriate, 1=very inappropriate and 5=very accurate, 1=very inaccurate). As the MARS instrument was segmented into 59 parts, each expert expressed a total of 118 ratings.

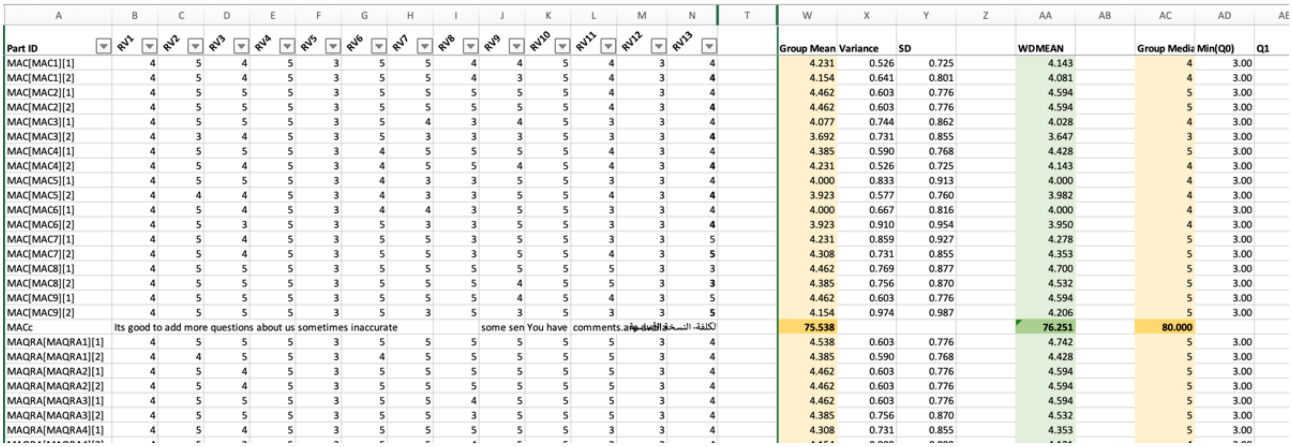
Out of the 19 available experts, 14 experts (14/19, 74% response rate) provided editing suggestions and completed the Web-based form evaluating the appropriateness and accuracy of the translated parts. An analysis of the Excel "comment dashboard" showed that experts provided a total of 287 editing suggestions for the MARS. In all, 3 reviewers provided editing suggestions for more than 50% of the MARS parts; 5 reviewers provided suggestions for more than 30%, and 6 reviewers provided suggestions for less than 30%. The parts that received the most editing suggestions (ie, from 10 to 14 reviewers) were the "Theoretical background/Strategies" and the "Technical aspects of app" in the "App Classification" section, followed by MARS item number 1, that is, "Entertainment" (*Is the app fun/entertaining to use? Does it use any strategies to increase engagement through entertainment, for example, through gamification?*), the description of Section A, that is, "Engagement" *(Engagement—fun, interesting, customizable, interactive—for example, sends alerts, messages, reminders, feedback, and enables sharing—and well targeted to audience*), and MARS item number 15 (*Quality of information: Is app content correct, well written, and relevant to the goal/topic of the app?*).

The research team created a matrix in Excel to track all comments and editing suggestions for each part of the translation. Each part was represented in rows, and the reviewers' comments were organized in columns. This "comment tracking dashboard" (Figure 2) was used to visually compare and contrast the comments received from the reviewers, which were color coded to simplify the reviewing process.

We created a similar matrix in Excel to calculate the level of agreement among experts. The "Interrater agreement (IRA) dashboard" (Figure 3) was used to calculate variance, means, and medians used to establish interrater agreement (IRA) according to the three families of indices: James et al's $r_{WG(J)}$ [47,48] (based on multiple null distributions [49]); Brown and Hauenstein's $a_{WG(J)}$ [50]; and the adjusted average deviation index $A_{DMJ(adj)}$ [51]. IRA was established through pragmatic and theoretical cutoff points, such as for the $r_{WG(J)}$: no agreement (<0.29), weak (0.30-0.49), moderate (0.50-0.69), strong (0.70-0.89), and very strong (>0.90) [52,53]; $a_{WG(J)}$: not acceptable (<0.59), weak (0.60-0.69), moderate (0.70-0.79), and strong agreement (>0.80) [50]; and $A_{DMJ(adj)}$: agreement above 0.80 [51].

Figure 2. Comment tracking dashboard.

Figure 3. Interrater agreement dashboard.



Part ID	RV1	RV2	RV3	RV4	RV5	RV6	RV7	RV8	RV9	RV10	RV11	RV12	RV13	Group Mean	Variance	SD	WDMEAN	Group Media	Min(Q0)	Q1
MAC[MAC1][1]	4	5	4	5	3	5	5	4	4	5	4	3	4	4.231	0.526	0.725	4.143	4	3.00	
MAC[MAC1][2]	4	5	4	5	3	5	5	4	3	5	4	3	4	4.154	0.641	0.801	4.081	4	3.00	
MAC[MAC2][1]	4	5	5	5	3	5	5	5	5	5	4	3	4	4.462	0.603	0.776	4.594	5	3.00	
MAC[MAC2][2]	4	5	5	5	3	5	5	5	5	5	4	3	4	4.462	0.603	0.776	4.594	5	3.00	
MAC[MAC3][1]	4	5	5	5	3	5	4	3	4	5	3	3	4	4.077	0.744	0.862	4.028	4	3.00	
MAC[MAC3][2]	4	3	4	5	3	5	3	3	3	5	3	3	4	3.692	0.731	0.855	3.647	3	3.00	
MAC[MAC4][1]	4	5	5	5	3	4	5	5	5	5	4	3	4	4.385	0.590	0.768	4.428	5	3.00	
MAC[MAC4][2]	4	5	4	5	3	4	5	5	4	5	4	3	4	4.231	0.526	0.725	4.143	4	3.00	
MAC[MAC5][1]	4	5	5	5	3	4	3	3	5	5	3	3	4	4.000	0.833	0.913	4.000	4	3.00	
MAC[MAC5][2]	4	4	4	5	3	4	3	3	5	5	4	3	4	3.923	0.577	0.760	3.982	4	3.00	
MAC[MAC6][1]	4	5	4	5	3	4	4	3	5	5	3	3	4	4.000	0.667	0.816	4.000	4	3.00	
MAC[MAC6][2]	4	5	3	5	3	5	3	3	5	5	3	3	4	3.923	0.910	0.954	3.950	4	3.00	
MAC[MAC7][1]	4	5	4	5	3	5	5	3	5	5	3	3	5	4.231	0.859	0.927	4.278	5	3.00	
MAC[MAC7][2]	4	5	4	5	3	5	5	3	5	5	4	3	5	4.308	0.731	0.855	4.353	5	3.00	
MAC[MAC8][1]	4	5	5	5	3	5	5	5	5	5	5	3	3	4.462	0.769	0.877	4.700	5	3.00	
MAC[MAC8][2]	4	5	5	5	3	5	5	5	4	5	5	3	3	4.385	0.756	0.870	4.532	5	3.00	
MAC[MAC9][1]	4	5	5	5	3	5	5	5	4	5	4	3	5	4.462	0.603	0.776	4.594	5	3.00	
MAC[MAC9][2]	4	5	5	5	3	5	3	5	3	5	3	3	5	4.154	0.974	0.987	4.206	5	3.00	
MACc	Its good to add more questions about us sometimes inaccurate								some sen	You have	comments are dwell			75.538			76.251	80.000		
MAQRA[MAQRA1][1]	4	5	5	5	3	5	5	5	5	5	5	3	4	4.538	0.603	0.776	4.742	5	3.00	
MAQRA[MAQRA1][2]	4	4	5	5	3	4	5	5	5	5	5	3	4	4.385	0.590	0.768	4.428	5	3.00	
MAQRA[MAQRA2][1]	4	5	4	5	3	5	5	5	5	5	5	3	4	4.462	0.603	0.776	4.594	5	3.00	
MAQRA[MAQRA2][2]	4	5	4	5	3	5	5	5	5	5	5	3	4	4.462	0.603	0.776	4.594	5	3.00	
MAQRA[MAQRA3][1]	4	5	5	5	3	5	5	4	5	5	5	3	4	4.462	0.603	0.776	4.594	5	3.00	
MAQRA[MAQRA3][2]	4	5	5	5	3	5	5	3	5	5	5	3	4	4.385	0.756	0.870	4.532	5	3.00	
MAQRA[MAQRA4][1]	4	5	4	5	3	5	5	5	5	5	5	3	4	4.308	0.731	0.855	4.353	5	3.00	

The “IRA dashboard” showed that the 14 experts rated the translation as highly appropriate (mean 4.37, SD 0.16; range 4.00-4.69) and accurate (mean 4.26, SD 0.20; range 3.62-4.69). The level of agreement was acceptable for most items, except for the “Interactivity” item (the fourth item of the domain “Engagement”). The level of agreement for the accuracy was not acceptable only for two parts: the “Theoretical background/Strategies” and the “Technical aspects of app” in the “App Classification” section.

The research team also compiled a Word document including all editing suggestions and comments and printed out the Excel “comment matrix” to easily visualize the suggestions. The research team met and discussed each comment, spending more than 8 hours reviewing the editing suggestions for each part of the MARS tool. The most debated parts were those including technical terms such as “goal setting” and “mindfulness” or “wellness,” which did not find an established equivalent term in Arabic. Notable changes from the original MARS included the removal of context-specific references that were not relevant to the Arab world, such as research funding sources provided in MARS item number 18 (ie, “Australian Research Council and National Health and Medical Research Council”). Minor editing was done in the response options for item number 2 of “Subjective Quality” (“How many times do you think you would use this app in the next 12 months if it were relevant to you?”): the anchor texts were changed to 11-50 to avoid overlap with the third option choice (3-10).

After the revisions were completed, the research team shared the edited Word document on Google Docs with the same pool of reviewers who participated in the first round, who were invited to comment by email. After 12 days, 5 experts provided 107 additional editing suggestions. For the second round, the research team did not collect quantitative measures to reduce the burden on the reviewers, as most of the editing work had already been done. The research team met once again to address (accept or reject) all comments and finalized the document.

The final version of the document was sent to a second professional translator, who was not involved in the process and had not seen the original MARS tool. The developer of the MARS approved the back-translation of the MARS-Ar. This document was used in the validation study (further described below).

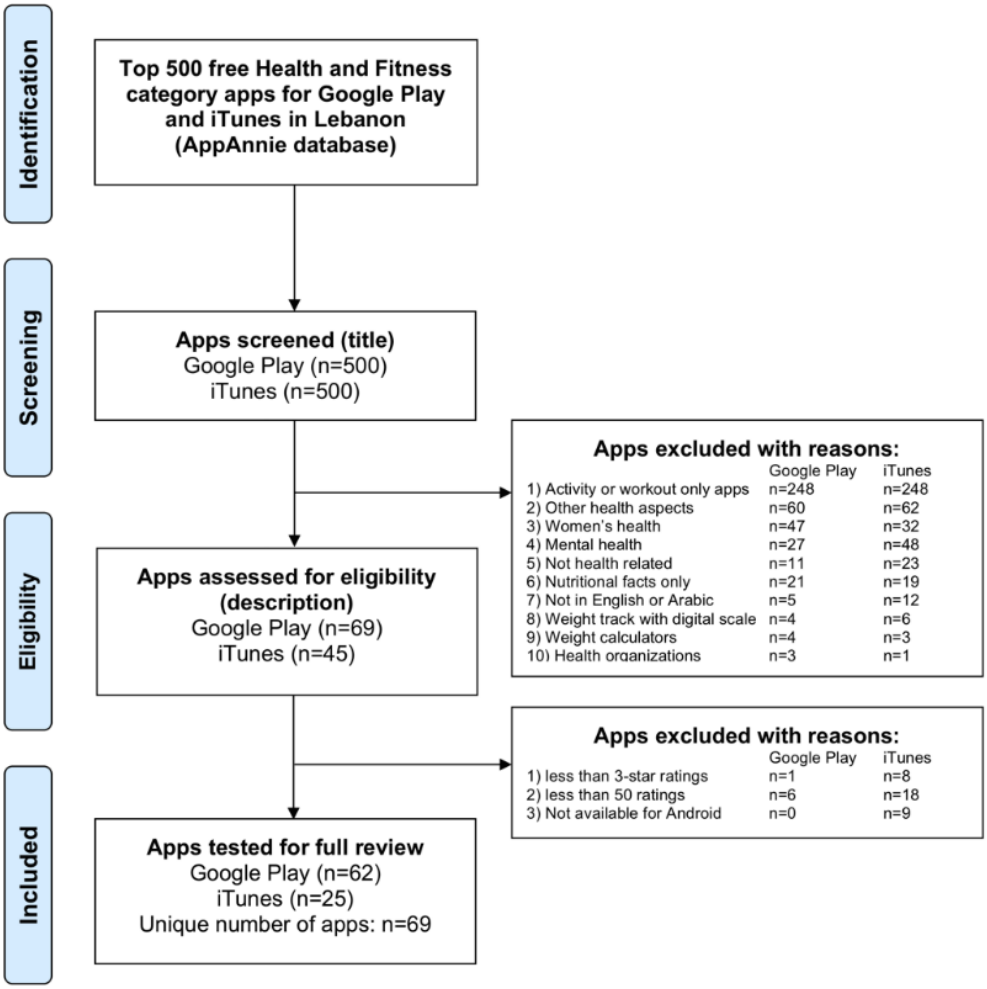
During the validation phase, one of the reviewers suggested some minor edits in the description of the “App Quality Ratings” part, in the description of the “Engagement” section, in the definition of “Target group” (item 5), in the description of the “Functionality” section, and in the items “Gestural design” (item 9) and “Graphics” (item 11). The research team approved the changes by circular vote. The final version of the MARS was then resent to the back-translator for verification. The final version of the MARS-Ar is available in Multimedia Appendix 1.

Phases 3 and 4: Piloting and Validation

App Selection Process

The research team identified the set of apps to be used in the piloting and validation phases of the study using the AppAnnie database (appannie.com), which provides updated rankings and mobile market data for both Android and iOS stores, under the section “App Store Rankings,” available after registering for free. On July 31, 2019, one researcher (MB) navigated the “Top Charts” section of the database, under the Google Play store, and filtered the country (Lebanon) and category (Health and Fitness) and selected the tab “Free” apps, extracting the titles and links to AppAnnie pages of 500 apps. These apps are listed under “free,” but in most cases, they operate under the “freemium” concept, with subscription fees used to remove ads and unlock complete features [54]. The researcher repeated the same procedure for the iOS store, as the apps’ rankings are quite different from the Google Play store, resulting in a second list of 500 apps. Links to AppAnnie’s webpages and titles of each app were imported in an Excel spreadsheet, to be screened for inclusion. The same researcher screened the lists and excluded irrelevant apps; a second researcher (NA) verified the selection. Any disagreement was discussed until consensus was reached. Of the total 1000 apps in both the Google Play store (Android) and the App Store (iOS), 431 and 455 apps were respectively excluded as they were not relevant (reasons for the exclusion are provided in the flowchart in Figure 4).

Figure 4. App selection process.



For the remaining 69 and 45 apps, the researchers extracted the following information from the AppAnnie's database: ranking in the Health and Fitness category of the respective store (Google Play or App Store), number of ratings, average 5-star rating, date of first release, date of last update, number of installs category, and price (for monthly subscription or yearly subscription). The dates of the first release and last update were used to calculate the "app age" in years.

On the basis of the number of ratings and average rating, 7 and 20 apps were excluded from Google Play and App Store lists, respectively, as they did not receive at least three stars or were not rated by at least 50 people. The researchers created a combined database of 78 unique apps that were available from either Google Play or App Store lists. Of these, nine apps were excluded as they were available only on the App Store list. The resulting 69 apps were used to validate the MARS-Ar tool, as reviewers owned only Android phones. Although there might be slight differences in the apps across iOS and Android operating systems, we have already established that these differences are not substantial [38].

The research team decided that the number of apps was sufficient to have reasonable empirical assurance and reliability, based on the intraclass correlation coefficients (ICCs), as reported in the source study [31], used in the Italian translation study [40], and on the basis of formulas described in the study by Zou [55]. For the Italian translation, Domnich et al [40] calculated a minimum sample size of 41 apps for two raters to achieve an assurance probability of 0.15 and an empirical assurance of 90%.

Rater Training

Two researchers (NA and TA), fluent in both Arabic and English and with a background in pharmacy, public health, and nutrition, completed independent evaluations of the selected apps. One of the two researchers was based in Jordan and was familiar with the MARS, as the researcher had previously used it. The second researcher was based in Lebanon. Both researchers were instructed to view the "MARS training video" in English (about 37 min, available on YouTube upon request from the author of the MARS). Thereafter, they were instructed to download each app on their phones (F1 Plus x9009 and Samsung S7 Edge, both with Android 5.1) and use them for at least 10 min, reporting any incompatibility issues, if they arose. Once the app was thoroughly tested, they individually and independently completed a Web-based form containing the MARS-Ar, available on LimeSurvey. After they completed the review of the apps in Arabic, they received a link to complete a form containing the original MARS in English, to establish a minimum criterion of validity with a validated "gold standard" instrument. The reviewers did not have access to the information

related to the apps so that users' ratings or reviews could not influence their evaluations.

Piloting and Evaluation

The 2 raters completed a calibration exercise using the first 10 apps in the list to ensure that both understood the meaning of all terminology correctly and that they could carefully review and discuss any points of difference in their ratings. We calculated interrater reliability using ICCs, based on a two-way mixed effect model in which people effects are random and measures effects are fixed, based on the example of previous MARS translation studies [40,42]. Reliability was interpreted as excellent (ICCs≥0.90), good (ICCs: 0.76-0.89), moderate (ICCs: 0.51-0.75), and poor (ICCs≤0.50). The ICC based on the ratings of the first 10 apps (23 items×10=230 decisions per rater) was moderate (ICC=0.714, 95% CI 0.619-0.785). The two reviewers met with the first author to discuss every rating that varied by 2 points or more. During the meeting, both raters aligned their rating approaches and confirmed their correct understanding of all MARS-Ar terminology. It was deemed that no further amendments to the scale were necessary. Finally, the two raters independently revised their responses and completed the evaluation of the remaining 59 apps on the list.

Analyses: Reliability and Internal Consistency

To verify whether the two raters provided comparable results among all the tested apps so that ratings could be aggregated, we assessed interrater reliability through ICCs, as described above. Once interrater reliability was ascertained, the individual ratings for each item of the MARS-Ar and original MARS were averaged. The resulting items were used to calculate the respective subdomain scales of engagement, functionality, aesthetics, information quality, and subjective quality. A total app quality score was calculated as the average of engagement, functionality, aesthetics, and information quality.

We also assessed internal consistency as a measure of scale reliability for the items pertaining to the same subdomain of the MARS, as reported in the original MARS study [31]. We used Cronbach alpha indices, interpreted as excellent (≥.90), good (.80-.89), acceptable (.70-.79), questionable (.60-.69), poor (.50-.59), and unacceptable (<.50).

As an indicator of validity, we used Pearson correlations between each subdomain score of the MARS-Ar and the MARS equivalent. In addition, we correlated the total MARS-Ar score, the total subjective quality score, and the subjective quality item number 4 (5-star rating) with the 5-star ratings from the app store to understand the extent to which reviewers' opinions about app quality were aligned with the users' opinions. A cutoff point of $r>0.80$ was deemed a sufficient indication of the validity of the MARS-Ar instrument.

All statistical tests were conducted using SPSS v21 [56] for macOS (Apple Inc, Cupertino, California).

Results

Evaluated Apps

The two reviewers completed the evaluation of 67 out of 69 selected apps, using MARS-Ar, and 66 apps, using the MARS English version. One app was incompatible with both test devices, and 2 apps were not working on one of the two devices used. Another app became unavailable for one device, as it was removed from the Google Play store when one of the reviewers completed the MARS-English form. The dataset of the tested apps, with statistics about their ranking, ratings, and age (since their first development), is available in Multimedia Appendix 2 (Excel file).

Interrater Reliability

The ICC based on the ratings for the full set of apps used in the MARS-Ar evaluation (23×67=1541 decisions per rater) was good (ICC=0.836, 95% CI 0.817-0.853). Similarly, the ICC for the English version (23×66=1518 decisions per rater) was also good (ICC=0.838, 95% CI 0.819-0.855).

Internal Consistency

Table 1 shows the overall descriptive statistics for both MARS-Ar and MARS English. All domains of MARS-Ar and original MARS showed good internal consistency. For MARS-Ar, internal consistency was good for engagement (Cronbach alpha=.96) and aesthetics (alpha=.94), good for information quality (alpha=.81), and acceptable for functionality (alpha=.71). Similar indices were also reported for the original MARS.

Overall, the tested set of weight management apps had high functionality and aesthetic scores but low engagement, information quality, and subjective quality scores.

Table 1. Summary of Mobile App Rating Scale in Arabic and Mobile App Rating Scale-English items and subdomains means, SDs, and Cronbach alpha coefficients.

Mobile App Rating Scale domains and subdomains	Mobile App Rating Scale in Arabic		Mobile App Rating Scale in English	
	Mean (SD)	Alpha	Mean (SD)	Alpha
Engagement	2.94 (0.99)	.95	3.12 (0.93)	.95
A1: Entertainment	2.69 (1.01)		2.78 (0.93)	
A2: Interest	2.87 (1.15)		3.21 (1.05)	
A3: Customization	2.69 (1.29)		2.86 (1.24)	
A4: Interactivity	2.66 (1.23)		2.85 (1.14)	
A5: Target group	3.78 (0.62)		3.89 (0.64)	
Functionality	4.11 (0.38)	.72	4.12 (0.32)	.73
B1: Performance	3.91 (0.71)		4.00 (0.53)	
B2: Ease of use	4.18 (0.37)		4.20 (0.30)	
B3: Navigation	4.22 (0.42)		4.10 (0.44)	
B4: Gestural design	4.13 (0.49)		4.17 (0.42)	
Aesthetics	3.14 (0.87)	.94	3.16 (0.72)	.96
C1: Layout	3.65 (0.74)		3.55 (0.71)	
C2: Graphics	3.00 (1.00)		2.98 (0.76)	
C3: Visual appeal	2.78 (1.00)		2.95 (0.78)	
Information quality	2.53 (0.73)	.81	2.59 (0.68)	.82
D1: Accuracy of app description	3.77 (0.64)		3.95 (0.53)	
D2: Goals	3.29 (0.99)		3.30 (0.86)	
D3: Quality of information	3.10 (0.90)		3.13 (0.75)	
D4: Quantity of information	2.51 (0.98)		2.80 (0.77)	
D5: Visual information	2.81 (1.86)		2.64 (1.86)	
D6: Credibility	1.99 (0.63)		1.90 (0.48)	
D7: Evidence base	0.27 (0.96)		0.42 (0.98)	
Subjective quality	2.21 (0.97)	.97	2.09 (0.79)	.95
SQ1: Would you recommend it?	2.34 (1.07)		2.11 (0.85)	
SQ2: How many times would you use it?	1.96 (1.03)		1.84 (0.87)	
SQ3: Would you pay for it?	1.75 (0.91)		1.67 (0.73)	
SQ4: 5-star rating	2.81 (1.07)		2.74 (0.90)	
Total app quality	3.18 (0.69)	—[a]	3.24 (0.61)	—

[a]Chronbach alpha for total app quality is not computed.

Mobile App Rating Scale in Arabic Validity

The correlations between MARS-Ar and original MARS and among each domain are presented diagonally in Table 2. The correlations among the domains of engagement, functionality, aesthetics, information quality, total app quality, and subjective quality are presented in the upper off-diagonal (for Arabic) and lower off-diagonal (for English).

The correlations between MARS-Ar and MARS-English were all significant at $P<.001$. The lowest was found in the domain of functionality ($r=0.685$), followed by aesthetics ($r=0.827$), information quality ($r=0.854$), engagement ($r=0.894$), and total app quality ($r=0.897$). Subjective quality scores and the item number 4 (5-star rating) were also highly correlated ($r=0.820$).

The 5-star rating from the app stores was not significantly associated with any app quality subdomain, total app quality, subjective quality, or MARS 5-star rating, neither in the Arabic nor in the English version.

Table 2. Correlations between Mobile App Rating Scale in Arabic and Mobile App Rating Scale-English domains and total app quality.

Mobile App Rating Scale in Arabic	Mobile App Rating Scale in English							5-star rating
	A	B	C	D	App quality	E	E4	
Engagement (A)	0.89[a,b]	0.64[a,c]	0.90[a,c]	0.92[a,c]	0.97[a,c]	0.90[a,c]	0.88[a,c]	−0.04[c]
Functionality (B)	0.61[a,d]	0.69[a,b]	0.70[a,c]	0.61[a,c]	0.75[a,c]	0.70[a,c]	0.48[a,c]	−0.03[c]
Aesthetics (C)	0.89[a,d]	0.68[a,d]	0.83[a,b]	0.86[a,c]	0.96[a,c]	0.91[a,c]	0.78[a,c]	0.03[c]
Information Quality (D)	0.92[a,d]	0.55[a,d]	0.81[a,d]	0.85[a,b]	0.95[a,c]	0.84[a,c]	0.80[a,c]	−0.14[c]
App quality score (average A-D)	0.96[a,d]	0.72[a,d]	0.95[a,d]	0.93[a,d]	0.90[a,b]	0.92[a,c]	0.84[a,c]	−0.05[c]
Subjective quality score (E)	0.88[a,d]	0.67[a,d]	0.82[a,d]	0.87[a,d]	0.90[a,d]	0.82[a,b]	0.78[a,c]	−0.07[c]
Subjective quality item number 4: 5-star rating (E4)	0.83[a,d]	0.62[a,d]	0.80[a,d]	0.78[a,d]	0.85[a,d]	0.83[a,d]	0.82[a,b]	0.00[c]
5-star rating (app stores)	−0.05[d]	0.04[d]	0.03[d]	−0.11[d]	−0.04[d]	−0.07[d]	−0.09[d]	1.00[b]

[a]$P<.001$.

[b]The diagonal shows the correlations between the same constructs of the MARS English and Arabic.

[c]In the upper diagonal section of the table: correlations among Mobile App Rating Scale subdomains, total app quality, and subjective quality (Mobile App Rating Scale in Arabic).

[d]In the lower diagonal section of the table: correlations among Mobile App Rating Scale subdomains (English).

Discussion

Principal Findings

This study aimed at translating and adapting MARS-Ar and at validating this scale with a set of popular health and fitness apps promoting weight management. The translation process demonstrated the importance of involving expert translators with interest and experience in translating technology-related documents. English-Arabic translation is not an easy task, as the language has many different regional varieties that make it difficult to find words that are common to the Modern Standard Arabic (MSA) dictionary [57]. In the literature related to English-Arabic translations, it is common to find reports of challenges related to the nonequivalence of words and sentence structures between the two languages [58], which occurs when translating colloquial or legal documents [59]. It was also important to involve experts from different countries of the Arab world, who provided valuable feedback and suggestions for improvement, as there are significant differences between the MSA and the many regional varieties (eg, Levantine Arabic vs Saudi or Gulf-countries or the Maghreb), with a plethora of dialects and different spoken expressions [60,61]. We found it challenging to find accurate translations of some technical terms and concepts referring to MARS domains, such as “Interactivity” or “Engagement,” which was also the case for some general terms, such as “goal setting” and “mindfulness” or “wellness,” usually used in disciplines such as Psychology and Health Sciences, usually taught in English; hence, the translations in Arabic were not easy to find.

After two rounds of review and additional feedback collected during the validation phase, we are confident to have a good instrument that Arabic-speaking researchers and experts can use to evaluate app quality in their native language. It is essential that Arabic-speaking researchers or professionals interested in evaluating apps establish a good and acceptable interrater reliability level before evaluating the full set of apps (ie, ICC above 0.70), as recommended in the MARS-German validation study [42]. A training video, similar to the one for MARS, will be developed so that the interpretation of terminology across researchers of different backgrounds and countries can be kept consistent.

This study’s results show that MARS-Ar is a reliable and valid instrument that trained “experts” can use to assess the quality of health apps. From a quantitative standpoint, there were no substantial differences in the reliabilities between the MARS-Ar and the original MARS in English. All MARS-Ar subdomains and individual quality items achieved appropriate internal consistency, comparable with the source study [31] and comparable with those reported in Italian [40] and German [42] validation studies. Similar to the German and Italian validation studies, the correlations between each subdomain of MARS-Ar and the original counterparts were also significant and extensive in size, indicating that the instrument tends to be valid.

In this study, we also found that the app quality ratings, according to experts, are not associated with the 5-star ratings reported in the app stores. These findings are consistent with another similar app review comparing expert ratings with the app stores [38] and with the MARS-German validation study [42]. App quality appears to be a complicated concept, which goes beyond a 5-star rating, as used in app stores. These ratings are not necessarily linked to the quality of health apps [62], as they can be inflated by developers [63]. With a sizeable and significant turnover of health apps [23], end users tend to rely on quick and available information to determine whether an app is worth downloading. MARS, as it is short and easy to understand and apply, could become the standard for app quality evaluation and provide researchers and end users with comparable dimensions across app domains.

With more versions of MARS available—Italian [40], Spanish [41], German [42], and now Arabic—it will be possible to

complete cross-cultural app evaluations and develop a joint research database of app evaluations, which could be made accessible to end users. Future studies should aim at involving end users to compare the ratings, for example, using the ratings between the uMARS and MARS versions.

The proposed project has a multifold impact. First, it provides Arab-speaking researchers and public health professionals, operating in the MENA region and elsewhere, with a culturally adapted and validated tool that could be used for developing new and evaluating existing apps. Second, this study will test whether MARS-Ar and uMARS in Arabic could be used to reliably evaluate the quality of apps for the prevention and treatment of obesity and related NCDs. Third, it can fulfill the needs of millions of people living in the region, who might be interested in knowing which apps could be trusted to prevent or better manage these conditions. Once the validation of the tool has been established, the researchers will maintain a database of app evaluations, thereby increasing the applicability and comparability of the results across multiple apps targeting the same public health issues.

Limitations

Despite its strengths, this study has some limitations to be acknowledged. First, the validity of the MARS-Ar instrument was established by comparing the scales in Arabic with their equivalents in the original MARS instrument, which the same raters completed in English. Future studies may compare MARS to other instruments of app quality [23,30], even though they might not be equivalent. We tested MARS-Ar with a set of apps for weight management; therefore, future studies need to test whether this instrument could also apply to health apps of different domains.

Conclusions

This study shows that MARS-Ar is a valid instrument, which can be used to assess app quality among trained Arabic-speaking users of health and fitness apps. Researchers and public health professionals in the Arab world can use the overall MARS score and its subscales to reliably evaluate the quality of weight management apps. Further studies are needed to test the instrument on health apps focusing on different health domains that are covered in health and fitness apps, such as mindfulness/anxiety prevention or sexual and reproductive health.

Acknowledgments

The authors wish to thank all the experts who contributed to the translation and contextualization of the MARS, in particular: Ms Reem Hoteit, Ms Farah Demachkieh, Ms Cosette Fakih El Khoury, Ms Tatyana Yousef El-Kour, Mr Amjad Hiary, Mr Mohamed A Abdel-Baqy. The authors also wish to thank the translators of the instruments. This study was supported by the University Research Board of the American University of Beirut, through a grant awarded to the first author (ref: 103608). The views expressed in this paper are those of the authors and do not necessarily reflect the funding agency or their institution

Authors' Contributions

MB designed and supervised the implementation of the study, coordinated the research activities, performed the analyses, drafted the manuscript, and incorporated all feedback from the coauthors. NA provided intellectual input to the study design, assisted in the implementation of the study, performed the app reviews, provided input on the translation of the instrument, and edited and provided feedback on the different versions of the manuscript. LG, EF, and TG provided intellectual input to the design of the study, assisted in the conduction of the study, provided input on the translation of the instrument, and edited and provided feedback on different versions of the manuscript. TA performed the app reviews and provided input on the translation of the instrument. SS provided intellectual input to the design of the study and on the methodology, and edited and provided feedback on different versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Arabic version of the Mobile App Rating Scale.
[DOCX File , 37 KB - mhealth_v8i3e16956_app1.docx]

Multimedia Appendix 2
Excel spreadsheet including the apps used in the validation study.
[XLSX File (Microsoft Excel File), 24 KB - mhealth_v8i3e16956_app2.xlsx]

References

1. World Health Organization. Geneva, Switzerland: World Health Organization; 2013. Global Action Plan for the Prevention and Control of NCDs 2013-2020 URL: http://www.who.int/nmh/publications/ncd_action_plan/en/ [accessed 2020-02-17]

2. Rahim HF, Sibai A, Khader Y, Hwalla N, Fadhil I, Alsiyabi H, et al. Non-communicable diseases in the Arab world. Lancet 2014 Jan 25;383(9914):356-367. [doi: 10.1016/S0140-6736(13)62383-1] [Medline: 24452044]
3. Musaiger A. Overweight and obesity in eastern mediterranean region: prevalence and possible causes. J Obes 2011;2011:407237 [FREE Full text] [doi: 10.1155/2011/407237] [Medline: 21941635]
4. Musaiger AO, Al-Mannai M, Al-Lalla O, Saghir S, Halahleh I, Benhamed MM, et al. Obesity among adolescents in five Arab countries; relative to gender and age. Nutr Hosp 2013 Nov 1;28(6):1922-1925 [FREE Full text] [doi: 10.3305/nutrhosp.v28in06.6412] [Medline: 24506370]
5. Badran M, Laher I. Type II Diabetes Mellitus in Arabic-Speaking countries. Int J Endocrinol 2012;2012:902873 [FREE Full text] [doi: 10.1155/2012/902873] [Medline: 22851968]
6. Hyder AA, Wosu AC, Gibson DG, Labrique AB, Ali J, Pariyo GW. Noncommunicable disease risk factors and mobile phones: a proposed research agenda. J Med Internet Res 2017 May 5;19(5):e133 [FREE Full text] [doi: 10.2196/jmir.7246] [Medline: 28476722]
7. The Lancet. Does mobile health matter? Lancet 2017 Nov 18;390(10109):2216. [doi: 10.1016/S0140-6736(17)32899-4] [Medline: 29165258]
8. Royston G, Hagar C, Long L, McMahon D, Pakenham-Walsh N, Wadhwani N, mHIFA Working Group (Mobile Healthcare Information For All). Mobile health-care information for all: a global challenge. Lancet Glob Health 2015 Jul;3(7):e356-e357 [FREE Full text] [doi: 10.1016/S2214-109X(15)00054-6] [Medline: 26087979]
9. International Telecommunication Union, World Health Organization. Be He@lthy, Be Mobile. Annual Report 2016. Geneva, Switzerland: International Telecommunication Union; 2016.
10. Statista. 2018. mHealth (Mobile Health) Industry Market Size Projection From 2012 to 2020 (in Billion US Dollars) URL: https://www.statista.com/statistics/295771/mhealth-global-market-size/ [accessed 2019-11-12]
11. Semper HM, Povey R, Clark-Carter D. A systematic review of the effectiveness of smartphone applications that encourage dietary self-regulatory strategies for weight loss in overweight and obese adults. Obes Rev 2016 Sep;17(9):895-906. [doi: 10.1111/obr.12428] [Medline: 27192162]
12. Aguilar-Martínez A, Solé-Sedeño JM, Mancebo-Moreno G, Medina FX, Carreras-Collado R, Saigí-Rubió F. Use of mobile phones as a tool for weight loss: a systematic review. J Telemed Telecare 2014 Sep;20(6):339-349. [doi: 10.1177/1357633X14537777] [Medline: 24875928]
13. Bardus M, Smith JR, Samaha L, Abraham C. Mobile phone and web 2.0 technologies for weight management: a systematic scoping review. J Med Internet Res 2015 Nov 16;17(11):e259 [FREE Full text] [doi: 10.2196/jmir.5129] [Medline: 26573984]
14. Quelly SB, Norris AE, DiPietro JL. Impact of mobile apps to combat obesity in children and adolescents: a systematic literature review. J Spec Pediatr Nurs 2016 Jan;21(1):5-17. [doi: 10.1111/jspn.12134] [Medline: 26494019]
15. Riaz S, Sykes C. Are smartphone health applications effective in modifying obesity and smoking behaviours? A systematic review. Health Technol 2015;5(2):73-81. [doi: 10.1007/s12553-015-0104-4]
16. Mateo GF, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015 Nov 10;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
17. Bardus M, Smith JR, Samaha L, Abraham C. Mobile and Web 2.0 interventions for weight management: an overview of review evidence and its methodological quality. Eur J Public Health 2016 Aug;26(4):602-610 [FREE Full text] [doi: 10.1093/eurpub/ckw090] [Medline: 27335330]
18. Schippers M, Adam PC, Smolenski DJ, Wong HT, de Wit JB. A meta-analysis of overall effects of weight loss interventions delivered via mobile phones and effect size differences according to delivery mode, personal contact, and intervention intensity and duration. Obes Rev 2017 Apr;18(4):450-459. [doi: 10.1111/obr.12492] [Medline: 28187246]
19. Research 2 Guidance. 2017. mHealth Economics 2017 – Current Status and Future Trends in Mobile Health URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2019-11-12] [WebCite Cache ID 6wOhZ1qY6]
20. Research and Markets. 2017 Nov. Global Mobile Health Market Analysis & Trends - Industry Forecast to 2025 URL: https://www.researchandmarkets.com/research/kxvn3l/global_mobile [accessed 2019-11-12]
21. Mikulich M. Statista. 2019 Sep 16. mHealth - Statistics & Facts URL: https://www.statista.com/topics/2263/mhealth/ [accessed 2019-11-07]
22. Statista. Projected Value of the Mobile Apps Market in Middle East and North Africa From 2014 to 2019, by Region (in Million US Dollars) URL: https://www.statista.com/statistics/454971/mena-app-market-value-region/ [accessed 2012-11-12]
23. Larsen ME, Nicholas J, Christensen H. Quantifying app store dynamics: longitudinal tracking of mental health apps. JMIR Mhealth Uhealth 2016 Aug 9;4(3):e96 [FREE Full text] [doi: 10.2196/mhealth.6020] [Medline: 27507641]
24. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014 May 16;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
25. Alnasser AA, Alkhalifa AS, Sathiaseelan A, Marais D. What overweight women want from a weight loss app: a qualitative study on arabic women. JMIR Mhealth Uhealth 2015 May 20;3(2):e41 [FREE Full text] [doi: 10.2196/mhealth.4409] [Medline: 25993907]

26. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
27. Nikolaou CK, Lean MEJ. Mobile applications for obesity and weight management: current market characteristics. Int J Obes (Lond) 2017 Jan;41(1):200-202. [doi: 10.1038/ijo.2016.186] [Medline: 27780974]
28. Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]
29. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. Telemed J E Health 2015 Feb;21(2):97-104. [doi: 10.1089/tmj.2014.0088] [Medline: 25469795]
30. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: a systematic review of prevalent and innovative methods. Am J Prev Med 2016 Dec;51(6):1051-1059. [doi: 10.1016/j.amepre.2016.07.009] [Medline: 27659122]
31. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
32. Wilson H, Stoyanov SR, Gandabhai S, Baldwin A. The quality and accuracy of mobile apps to prevent driving after drinking alcohol. JMIR Mhealth Uhealth 2016 Aug 8;4(3):e98 [FREE Full text] [doi: 10.2196/mhealth.5961] [Medline: 27502956]
33. Furlong LM, Morris ME, Erickson S, Serry TA. Quality of mobile phone and tablet mobile apps for speech sound disorders: protocol for an evidence-based appraisal. JMIR Res Protoc 2016 Nov 29;5(4):e233 [FREE Full text] [doi: 10.2196/resprot.6505] [Medline: 27899341]
34. Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the Mobile Application Rating Scale (MARS). JMIR Mhealth Uhealth 2016 Jun 14;4(2):e74 [FREE Full text] [doi: 10.2196/mhealth.5882] [Medline: 27302310]
35. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
36. Zini F, Reinstadler M, Ricci F. Increasing quality of life awareness with life-logging. In: Giokas K, Bokor L, Hopfgartner F, editors. eHealth 360°. Cham: Springer; 2017:282-293.
37. Patel R, Sulzberger L, Li G, Mair J, Morley H, Shing MN, et al. Smartphone apps for weight loss and smoking cessation: quality ranking of 120 apps. N Z Med J 2015 Sep 4;128(1421):73-76. [Medline: 26370762]
38. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016 Mar 10;13:35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
39. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
40. Domnich A, Arata L, Amicizia D, Signori A, Patrick B, Stoyanov S, et al. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. BMC Med Inform Decis Mak 2016 Jul 7;16:83 [FREE Full text] [doi: 10.1186/s12911-016-0323-2] [Medline: 27387434]
41. Payo RM, Álvarez MM, Díaz MB, Izquierdo MC, Stoyanov S, Suárez EL. Spanish adaptation and validation of the Mobile Application Rating Scale questionnaire. Int J Med Inform 2019 Sep;129:95-99. [doi: 10.1016/j.ijmedinf.2019.06.005] [Medline: 31445295]
42. Messner E, Terhorst Y, Barke A, Baumeister H, Stoyanov S, Hides L, et al. Development and validation of the German version of the Mobile Application Rating Scale (MARS-G). JMIR Preprints 2019. [doi: 10.2196/14479]
43. EF Education First. Middle East URL: https://www.ef.com/epi/regions/middle-east/ [accessed 2018-02-18]
44. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. Qual Life Res 1998 May;7(4):323-335. [doi: 10.1023/a:1024985930536] [Medline: 9610216]
45. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. J Clin Epidemiol 2015 Apr;68(4):435-441. [doi: 10.1016/j.jclinepi.2014.11.021] [Medline: 25698408]
46. LimeSurvey. URL: https://www.limesurvey.org/ [accessed 2020-02-17]
47. James LR, Demaree RG, Wolf G. Estimating within-group interrater reliability with and without response bias. J Appl Psychol 1984;69(1):85-98. [doi: 10.1037/0021-9010.69.1.85]
48. James LR, Demaree RG, Wolf G. rwg: An assessment of within-group interrater agreement. J Appl Psychol 1993;78(2):306-309. [doi: 10.1037/0021-9010.78.2.306]
49. Biemann T, Cole MS, Voelpel S. Within-group agreement: On the use (and misuse) of rWG and rWG(J) in leadership research and some best practice guidelines. Leadersh Q 2012;23(1):66-80. [doi: 10.1016/j.leaqua.2011.11.006]
50. Brown RD, Hauenstein NM. Interrater agreement reconsidered: an alternative to the rwg indices. Organ Res Methods 2005;8(2):165-184. [doi: 10.1177/1094428105275376]
51. Lohse-Bossenz H, Kunina-Habenicht O, Kunter M. Estimating within-group agreement in small groups: a proposed adjustment for the average deviation index. Eur J Work Organ Psychol 2014;23(3):456-468. [doi: 10.1080/1359432X.2012.748189]

52. LeBreton JM, Senter JL. Answers to 20 questions about interrater reliability and interrater agreement. Organ Res Method 2008;11(4):815-852. [doi: 10.1177/1094428106296642]
53. O'Neill TA. An overview of interrater agreement on Likert scales for researchers and practitioners. Front Psychol 2017;8:777 [FREE Full text] [doi: 10.3389/fpsyg.2017.00777] [Medline: 28553257]
54. Kosner AW. Forbes. 2015 Jan 11. Apple App Store Revenue Surge And The Rise Of Freemium App Pricing URL: http://www.forbes.com/sites/anthonykosner/2015/01/11/apple-app-store-revenue-surge-and-the-rise-of-the-freemium/2/ [accessed 2015-04-09] [WebCite Cache ID 6Xf9egBFu]
55. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Stat Med 2012 Dec 20;31(29):3972-3981. [doi: 10.1002/sim.5466] [Medline: 22764084]
56. IBM. IBM SPSS Statistics Version 21.0 URL: https://www.ibm.com/support/pages/spss-statistics-210-available-download [accessed 2020-02-17]
57. Benzehra R. Issues and challenges for a modern English-Arabic dictionary. Dictionaries 2012;33(1):83-102. [doi: 10.1353/dic.2012.0008]
58. Chahrour O. Translation Journal. 2017 Apr. Cultural Difficulties in Translations from English Into Arabic URL: https://translationjournal.net/April-2017/cultural-difficulties-in-translations-from-english-into-arabic.html [accessed 2019-10-13]
59. Zidan AA. A Linguistic Analysis of Some Problems of Arabic-English Translation of Legal Texts, with Special Reference to Contracts. Cambridge: Cambridge Scholars Publishing; 2015.
60. Huddleston G. Language Connect. 2017 Feb 28. Challenges in translation - Arabic to English and other languages - Part two URL: https://www.languageconnect.net/blog/language-connect/challenges-in-translation-arabic-to-english-and-other-languages-part-two/ [accessed 2019-10-13]
61. Huddleston G. Language Connect. 2017 Feb 28. Challenges in Translation - Arabic to English and Other Languages: Part One URL: https://www.languageconnect.net/blog/language-connect/challenges-in-translation-arabic-to-english-and-other-languages-part-one/ [accessed 2019-10-13]
62. Plante T, O'Kelly AC, Macfarlane Z, Urrea B, Appel L, Miller III ER, et al. Trends in user ratings and reviews of a popular yet inaccurate blood pressure-measuring smartphone app. J Am Med Inform Assoc 2018 Aug 1;25(8):1074-1079 [FREE Full text] [doi: 10.1093/jamia/ocy060] [Medline: 29878236]
63. Hill S. Digital Trends. 2018 Mar 3. Can You Really Trust App Store Ratings? We Asked the Experts URL: https://www.digitaltrends.com/android/can-you-really-trust-app-store-ratings/ [accessed 2019-10-13]

Abbreviations

ICC: intraclass correlation coefficient
IRA: interrater agreement
MARS: Mobile App Rating Scale
MARS-Ar: Mobile App Rating Scale in Arabic
MENA: Middle East and North Africa
mHealth: mobile health
MSA: Modern Standard Arabic
NCD: noncommunicable disease
uMARS: user version of the Mobile App Rating Scale

Edited by G Eysenbach; submitted 07.11.19; peer-reviewed by A Domnich, RM Payo, C Reis; comments to author 28.11.19; revised version received 14.12.19; accepted 15.12.19; published 03.03.20.

Please cite as:
Bardus M, Awada N, Ghandour LA, Fares EJ, Gherbal T, Al-Zanati T, Stoyanov SR
The Arabic Version of the Mobile App Rating Scale: Development and Validation Study
JMIR Mhealth Uhealth 2020;8(3):e16956
URL: https://mhealth.jmir.org/2020/3/e16956
doi:10.2196/16956
PMID:32130183

Review

Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review

Madison Milne-Ives[1], BAS, MSc; Ching Lam[1], MEng; Caroline De Cock[1], BSc, MSc; Michelle Helena Van Velthoven[1], BSc, MSc, PhD; Edward Meinert[1,2], MA, MSc, MBA, MPA, PhD

[1]Digitally Enabled Preventative Health Research Group, Department of Paediatrics, University of Oxford, Oxford, United Kingdom

[2]Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

Corresponding Author:
Edward Meinert, MA, MSc, MBA, MPA, PhD
Digitally Enabled Preventative Health Research Group
Department of Paediatrics
University of Oxford
Headley Way, Headington
John Radcliffe Hospital
Oxford, OX3 9DU
United Kingdom
Phone: 44 7824446808
Email: e.meinert14@imperial.ac.uk

Abstract

Background: With a growing focus on patient interaction with health management, mobile apps are increasingly used to deliver behavioral health interventions. The large variation in these mobile health apps—their target patient group, health behavior, and behavioral change strategies—has resulted in a large but incohesive body of literature.

Objective: This systematic review aimed to assess the effectiveness of mobile apps in improving health behaviors and outcomes and to examine the inclusion and effectiveness of behavior change techniques (BCTs) in mobile health apps.

Methods: PubMed, EMBASE, CINAHL, and Web of Science were systematically searched for articles published between 2014 and 2019 that evaluated mobile apps for health behavior change. Two authors independently screened and selected studies according to the eligibility criteria. Data were extracted and the risk of bias was assessed by one reviewer and validated by a second reviewer.

Results: A total of 52 randomized controlled trials met the inclusion criteria and were included in the analysis—37 studies focused on physical activity, diet, or a combination of both, 11 on drug and alcohol use, and 4 on mental health. Participant perceptions were generally positive—only one app was rated as less helpful and satisfactory than the control—and the studies that measured engagement and usability found relatively high study completion rates (mean 83%; n=18, N=39) and ease-of-use ratings (3 significantly better than control, 9/15 rated >70%). However, there was little evidence of changed behavior or health outcomes.

Conclusions: There was no strong evidence in support of the effectiveness of mobile apps in improving health behaviors or outcomes because few studies found significant differences between the app and control groups. Further research is needed to identify the BCTs that are most effective at promoting behavior change. Improved reporting is necessary to accurately evaluate the mobile health app effectiveness and risk of bias.

(JMIR Mhealth Uhealth 2020;8(3):e17046) doi:10.2196/17046

KEYWORDS

telemedicine; evidence-based medicine; mobile health; digital health; mobile applications; app; cell phone; smartphone; mobile phone; health behavior; intervention; behavior change; systematic review

Introduction

Background

Engaging patients with health care is an important area of development in health care because it has the potential to reduce preventable deaths [1,2]. There is a huge range of digital health technologies that can deliver health care interventions, including apps, SMS texts, emails, internet, interactive chatbots, and voice agents [3-5]. Since the first iPhone was released in 2008, smartphone technology has become increasingly prevalent and capable, offering a promising means of delivering health care interventions to the general population. The large number of mobile health apps currently available for download is a testament to their popularity [6]. Many mobile phones now have the ability to passively collect a variety of health data—including physical activity, social interaction, sleep, and mobility patterns— and make inferences about mental and physical health [7,8]. Combining these capabilities with active user interaction allows mobile apps to deliver many different behavioral interventions, which can help users lead healthier lives and potentially reduce the likelihood of preventable health issues.

Mobile apps have been designed to target a wide variety of actions to prevent problems and maintain and improve patients' health [9]. There are five main types of health behaviors—physical activity, diet, drug use, alcohol use, and mental health [4]—but other actions such as the management of chronic conditions, medication adherence, doctor appointments, vaccinations, dental hygiene, sun protection, and sex safety can also be considered health behaviors [10,11]. However, mobile health apps' effectiveness has not been sufficiently established [4,12,13]. Many studies do not even report whether or not their mobile health behavior apps are based on behavioral theories. Although there is a debate on the role behavioral interventions can and should play in the population-level behavior, behavioral theory is agreed to be an important component of successful health-related behavioral interventions [14]. Further evaluation of the effectiveness of mobile health apps is needed to determine which apps are most useful and which behavioral change theories and techniques best promote positive behavior change, which in turn can guide future development. This is important because of the ubiquity of mobile health apps in society—if they are to fulfill their intended goal of improving health, they must be able to effectively improve and maintain positive health behaviors.

Many systematic reviews are currently examining these topics. However, with a few exceptions [12,15], most of these reviews are restricted in scope to specific types of health behaviors, patient groups, or combinations thereof. This has the advantage of being able to more directly compare the studies, and perform meta-analyses, if the studies are similar enough. The diversity of the mobile health app field makes it difficult to coalesce the results of studies into a coherent overview. A few systematic reviews have taken on this challenge. These examined studies focused on a variety of health behaviors, patient groups, or types of app intervention. Overall, they found that mobile health apps were effective in improving participants' health behaviors [12,15]. They also identified the most commonly used behavior change constructs: self-monitoring [13,15], cues to action, feedback, and social support [15] However, each of these systematic reviews has limitations. These limitations include the exclusive use of broad search terms, which likely missed many relevant mobile health app articles that used more specific key terms [12], and that the articles reviewed were primarily pilot studies with small sample sizes [15]. In addition, data for these systemic reviews were collected in 2014 [15], 2015 [13], and 2017 [12]. Given the rapid pace of technological development, a new systematic review is necessary to provide an accurate assessment of the effectiveness of the most recent mobile health app interventions.

Identifying the behavior change techniques (BCTs) that are most effective in fostering positive change is necessary to develop the most effective and engaging interventions to improve the participant health behavior. An update to and expansion of previous systematic reviews is needed to provide an overview of current mobile health app technology, the BCTs being used, and their effectiveness in changing behavior and participant health outcomes. New, innovative apps are continuously being developed and tested, and systematic reviews must keep pace so that overall trends in the features, theories, and effectiveness of these apps can be tracked and updated. To ensure that the mobile health apps that patients are using are achieving their promises of health behavior change, it is essential to have a clear understanding of what is currently being used and whether it is working. If barriers—in the apps' features or use—can be identified, app developers can use that information to design more effective interventions.

Objectives

The primary objectives of the review are to summarize the state of the field of mobile health apps for behavior change and evaluate their effectiveness. The wide variety of apps and health behaviors examined means that there are a wide range of outcomes examined to address three main research questions. First, what types of mobile health apps and BCTs are being used to support user engagement with their health behaviors? Second, how effective are mobile health apps in improving and maintaining positive health behavior changes? Finally, what are participant perceptions of the feasibility, functionality, and overall user experience of the mobile health apps they use? These are the key elements that are needed to comprehensively evaluate mobile health apps. This focus builds on previous systematic reviews and extends and updates the body of knowledge on current mobile health apps to inform further research and development.

Methods

Database Search

The methods are described in detail in a systematic review protocol that is registered with the International Prospective Register of Systematic Reviews (PROSPERO: CRD42019155604). The search strategy was developed using the Population, Intervention, Comparison, and Outcome (PICO) template and performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P Multimedia Appendix 1) [16]. MeSH terms and

keywords were extracted from a preliminary review of the literature, and the search strings and databases were decided in consultation with a medical librarian. The search was performed in four databases using the University of Oxford Search Oxford Libraries Online—PubMed, EMBASE, Cumulative Index to Nursing and Allied Health Literature, and Web of Science—with slightly adjusted search terms to fit the specific structure of each database. The search terms were grouped into four themes—mobile phones, mobile apps, health behaviors, and evaluation—that were joined with the structure: Mobile (MeSH OR Keywords) AND Applications (MeSH OR Keywords) AND Health Behavior (MeSH OR Keywords) AND Evaluation (MeSH OR Keywords). A complete list of the specific search terms and strings used for each database is provided in Multimedia Appendices 2 and 3. The search was completed on September 16, 2019.

Inclusion and Exclusion Criteria

Digital health technologies evolve rapidly, and this review was concerned with the current state of mobile health app technology [17]. Therefore, the search was limited to studies published between 2014 and 2019. This time frame provided an update to Payne et al's [15] systematic review that included studies published between 2007 and September 2014 and reflected the most recent behavioral and technological developments. Only studies published in English were included.

This systematic review had a broad scope for population and included any age, gender, country, or ethnicity. Therefore, study populations could be general or specific with regard to demographic variables. However, to keep the focus on the effectiveness of mobile health apps in the general population, subgroups such as pregnant women and patients with specific diseases (including HIV, posttraumatic stress disorder, alcoholism, and chronic depression) were excluded.

The intervention targeted in this review was mobile apps for health behavior change. Therefore, to be included, the main focus of the study needed to be the evaluation of a mobile app that helps users adopt, improve, or maintain positive health behaviors. Studies of mobile interventions that did not evaluate the app, were designed for use by health care professionals, focused on behavior change theory without reference to mobile apps, or focused on mobile phones or wearable technology but not apps—for instance, interventions based solely on SMS text messaging or emails—were excluded. Interventions that were primarily focused on mobile apps but involved wearable technology for data collection were included.

Initially, we intended to include all types of health behaviors. However, the number of studies that this would have entailed was unfeasible, and the health behaviors that were included in the final systematic review were limited to the five main categories established in the literature: drug use, alcohol use, diet, physical activity, and mental health. This excluded behaviors such as sun protection, sex safety, medication adherence, doctor's appointments, vaccinations, and self-management of chronic conditions.

Study design was not limited in the initial search to ensure that no relevant studies were missed, but only randomized controlled trials were included in the review. All types of comparators were included.

Outcomes Measured

The primary outcomes were participant health and behavior change to evaluate the apps' effectiveness. Secondary outcomes included the apps' features and their adoption of specific BCTs [18], as well as engagement and adherence rates, participant-reported experience, and feasibility and usability assessments.

Screening and Study Selection

All the articles identified from the database searches were stored in the citation management software to eliminate duplicates before the abstracts were screened by two independent reviewers. Disagreements were discussed until consensus on eligibility was reached. The full text was screened by one of the reviewers, and when the text did not meet the inclusion criteria, the second reviewer reviewed the article to assess eligibility to determine inclusion in the final set. Reasons for inclusion and exclusion were recorded at both the abstract and full-text screening stages.

To check if the search had missed any relevant articles, the full citation list was compared with the list of studies included in the two previous systematic reviews [12,15]. Of the 20 articles examined in Han and Lee's [12] review, 11 were already included in the citation list, and none of the other 9 were eligible for inclusion [12]. As Payne et al's [15] review finished data collection in 2014, only 8 of their 24 studies were within the time frame of this review [15]. Of those, 6 were already in the citation list, 1 had just been identified from the Han and Lee's review [12], and the other was added to the overall citation list but excluded from the final review because it was a treatment for major depressive disorder.

A total of 8 of the screened articles eligible for inclusion were abstracts from posters, conferences, or meetings and did not have full texts available. The authors of each were contacted to request a full text if it was available. At the time of writing, only one of the study authors had replied. A full text was sent but ended up not being relevant and was excluded from the final review because a mobile app was not the main focus of the study.

Data Extraction

Data were extracted by one reviewer and key data points from the studies that were specified in the protocol were recorded in a spreadsheet (see Table 1). This process was validated by the second reviewer, and disagreements were discussed with a third reviewer. The broad scope of the review meant that there were a wide variety of specific health and behavior change outcomes, so a meta-analysis could not be performed.

Table 1. Data that were extracted from the studies.

Article information	Data extracted
General study information	• Year of publication • Countries of study • Study setting (primary location of app use, if relevant) • Analyzed sample size • Sample demographics (including age, gender, and target population) • Intervention duration and follow-up periods
Behavioral intervention	• Target health behaviors and intervention focus • Theory the intervention is based on • Behavior change techniques (BCT Taxonomy v1 [18,19])
Mobile app technology	• Area of health care used in • Name of the app • Developers • Platform • Components and design features (eg, provision of feedback, notifications, and tracking)
Evaluation	• What outcomes were measured • Participant health outcomes • Behavior change outcomes • Participant engagement or adherence rates • Participant satisfaction • Feasibility and usability • Other key performance indicators reported

Risk of Bias Assessment

The study design was limited to randomized controlled trials, so we used the Cochrane Collaboration Risk of Bias tool to assess all of the included articles [20]. Specifically, this assessed the risk of bias in random sequence generation; allocation concealment; blinding of participants, personnel, and outcome assessors; incomplete outcome data; and selective outcome reporting. The risk of bias assessment was conducted by one reviewer and validated by the second reviewer, and disagreements were resolved by discussion.

Data Analysis and Synthesis

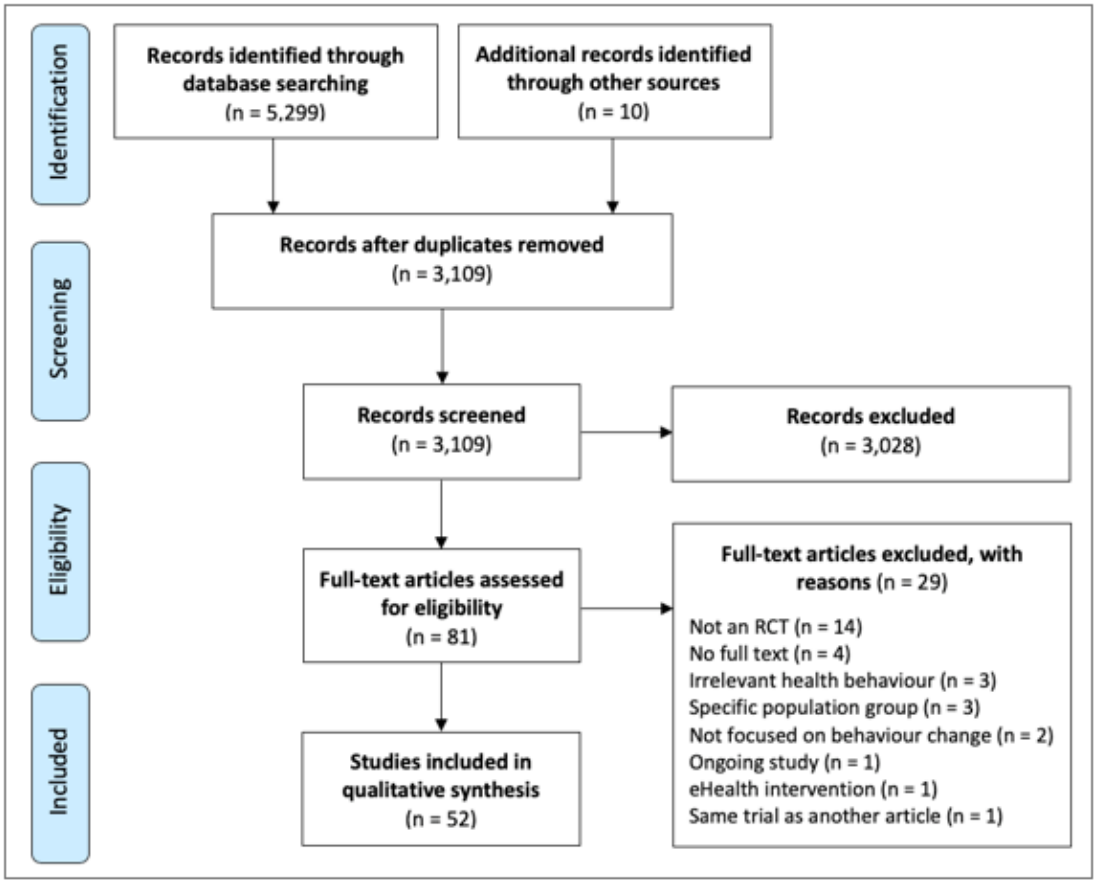
The variety of study aims, methods, and reported outcomes meant that a meta-analysis was unfeasible, so a narrative summary of the studies was prepared to draw conclusions about the apps' effectiveness, use of BCTs, acceptability, and usability. In this review, an outcome was only considered to have significant evidence supporting it if the app performed significantly better than a comparator or control. Outcomes that were significantly different over time but not between groups were coded as having some evidence supporting them. Outcomes that were not significantly different between groups, had no significant effect, or were significantly worse than the comparator were coded as having no evidence supporting them. Limitations and future directions for research and development were also identified.

Results

Included Studies

In total, we retrieved 5299 articles using the search terms in the four databases. Of these, 81 were selected for the full-text review, and 52 were selected for inclusion in the review. The reasons for exclusion in the full-text review stage are detailed in Figure 1. One article was excluded at the full-text stage for reporting the same trial as another article that was already included. This excluded article was more focused on the development of the app and did not provide any additional relevant data for extraction.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram. eHealth: electronic health; RCT: randomized controlled trial.



Study Characteristics

The characteristics of the 52 included studies are summarized in Multimedia Appendix 4. Of these 52, 71% (n=37) aimed to change dietary habits, physical activity, or both to reduce or prevent obesity and improve general health. The goal of 21% (11/52) of the studies was to reduce drug and alcohol use (9/52 were related to smoking cessation, and 2/52 aimed to reduce alcohol consumption). The last 8% (4/52) targeted behaviors aimed at improving mental health. More than half of the studies (28/52) had sample sizes of less than 100 participants, and slightly fewer than half (23/52) had a study duration less than 3 months.

Overall Effectiveness of Apps

One study was excluded from this part of the analysis because it only evaluated feasibility, usability, and participant perceptions [21]. Only about a quarter of the studies (12/51) found that the app had a significantly better effect (P<.05) on participant health or behavior change outcomes than control or comparator groups (see Table 2). These include significantly bigger increases for the app group in healthy food consumption (eg, 1-2 more daily servings of vegetables [22,23]), physical activity (eg, 1000-2000 more steps per day [24,25]), and mental health (large Cohen d effect sizes for mindfulness and self-compassion , 0.8 and 1.1, respectively [26]), as well as significantly bigger decreases in drug and alcohol use [27,28]. About 30% of the studies (16/51) found some evidence of effectiveness—whether there was a significant difference between the groups on some but not all of the outcomes, a significant difference over time but not between groups, or a significant improvement over the control only in a subgroup of the population. The remaining 45% of studies either found no significant difference between groups or effect on the primary outcome (22/51) or found that the app performed worse than the comparator (1/51). This analysis was done by coding the overall outcomes of each study, but the ratio is very similar if the overall effectiveness is coded once for each individual app instead.

Table 2. Summary of general evidence of effectiveness by study (N=51).

Evidence of effectivenes?	Physical activity (PA) [24,25,29-39], n	Diet [22,23,40-49], n	Diet and PA[a] [50-60], n	Mental health [26,61-63], n	Smoking cessation [28,64-71], n	Reduce alcohol [27,72], n	Total, n (%)
No	4	3	7	1	7	1	23 (45)
Some	4	5	4	2	1	0	16 (31)
Yes	5	4[b]	0	1	1	1	12 (23)
Total	13	12	11	4	9	2	51 (100)

[a]The studies in the diet and physical activity category reported on dietary and physical activity outcomes, whereas the studies in the previous 2 columns reported on either diet or physical activity.

[b]Two of these studies report on the same trial (one at 12 weeks and the other at the end of the 12-month trial) [22,47]. Both have been included in this table, but if one were excluded, there would be significant evidence for 22% (11/50) studies.

Participant Health Outcomes

A wide range of participant health outcomes were reported in 20 studies, including measures of weight change over time (weight, BMI, waist circumference, and body adiposity), mental well-being (depression, anxiety, life satisfaction, perceived stress, emotional regulation, etc), blood pressure, cardiovascular risk factors, and biomarkers such as blood lipids and urinary sodium. Overall, there was very little evidence that supported the effectiveness of mobile health apps to affect participant health outcomes (see Table 3). Over three-quarters (24/31) of the reported participant health outcomes were not significantly different between the intervention and control groups.

Table 3. Effectiveness of apps on participant health outcomes (N=31).

Participant health outcome	No evidence, n	Some evidence, n	Significant evidence, n (%)	Studies reporting outcome, n
Weight/BMI change [35,41-46,48,50,55,59,60]	10	1	1 (8)	12
Waist circumference/body adiposity [43,53,59]	2	1	0 (0)	3
Mental well-being (eg, depression, anxiety, perceived stress, life satisfaction, and mood) [26,50,61-63]	2	3	0 (0)	5
Blood pressure [35,41,46,51,54]	5	0	0 (0)	5
Cardiovascular risk factors [54,55]	2	0	0 (0)	2
Blood measures (eg, blood glucose and blood lipids) [41,54]	2	0	0 (0)	2
Urinary sodium [43,51]	1	1	0 (0)	2
Total	24	6	1 (3)	31

Behavior Change Outcomes

An even broader range of behavior change outcomes were reported in 44 of the 52 studies. Overall, there was not much significant evidence supporting the effectiveness of apps in changing the behavior outcomes (see Table 4). There were certain types of behavior that had stronger evidence of change than others. A total of 63% (5/8) of studies that examined healthy food choice behavior found that the app group improved significantly more than the control or comparator group, as did 43% (3/7) of the studies that reported step count and 100% (3/3) of the studies that aimed to reduce sedentary behavior. Physical activity and dietary habits were the target behavior areas with the highest percentage of significant evidence (32% each). However, there were only a few studies for each specific outcome, and altogether, the studies only found significant support for the effectiveness of mobile health apps in just a quarter (16/64) of the behavior change outcomes reported.

Table 4. Effectiveness of apps with respect to behavior change outcomes (N=44).

Target behavior and behavior change outcome[a]	No evidence, n	Some evidence, n	Significant evidence, n (%)	Total times outcome reported, n
Dietary habits				
Healthy food choices (including vegetable consumption and purchase of salt) [22,23,40,44,45,48,51,59]	2	1	5 (63)	8
Hunger [40]	1	0	0 (0)	1
Control (including cognitive restraint, self-efficacy, self-regulation, PBC[b], and avoiding uncontrolled eating) [40,41,45,46,49,60]	5	0	1 (17)	6
Dietary compliance (including goal setting and diet tracking) [42,45,46,52,58]	4	0	1 (20)	5
Energy/caloric intake [55,57]	1	1	0 (0)	2
Total (dietary habits)	13	2	7 (32)	22
Physical activity				
Physical activity (including moderate to vigorous physical activity) [25,30,33,37-39,52,55,58]	7	1	1 (11)	9
Walking/step count [24,25,29,33,35,57]	2	2	3 (43)	7
Reduce sedentary behavior [29,32,37]	0	0	3 (100)	3
Time to complete fitness test [31]	1	0	0 (0)	1
Attitudes to physical activity [34]	0	1	0 (0)	1
Control (including self-efficacy, PBC, and barriers) [25,34,39]	2	1	0 (0)	3
Self-monitoring [39]	0	0	1 (100)	1
Total (physical activity)	12	5	8 (32)	25
Reduce alcohol				
Change in weekly alcohol consumption [27]	1	0	0 (0)	1
Change in full Alcohol Use Disorders Identification Test score [27]	1	0	0 (0)	1
Number of alcohol consequences [72]	1	0	0 (0)	1
Maximum number of drinks at once [72]	1	0	0 (0)	1
Total (reduce alcohol)	4	0	0 (0)	4
Smoking cessation				
Continuous abstinence (including 7- and 30-day point prevalence abstinence) [28,64,67-71]	5	1	1 (14)	7
Quit rates [65,66]	2	0	0 (0)	2
Acceptance of cravings [65]	0	1	0 (0)	1
Readiness to quit (including motivation and quit attempts) [66,69,71]	3	0	0 (0)	3
Total (smoking cessation)	10	2	1 (8)	13
Total	39	9	16 (25)	64

[a]Many of the studies reported more than one behavior change outcome, and all distinct outcomes were recorded here, so there are more individual outcomes than the number of studies.

[b]PBC: perceived behavioral control.

Behavior Change Techniques and Theories

In the 52 studies, there were 50 unique apps tested (excluding basic or control versions of apps). Only a few of these studies explicitly reported the BCTs incorporated into the app, so BCTs were coded based on the study descriptions of the app features using the BCT Taxonomy v1 [18]. The taxonomy lists 93 BCTs, clustered into 16 groups. Collectively, the apps studied included 39 different BCTs from 12 different groups. Multimedia Appendix 5 reports the BCTs included in each app studied.

Only four BCTs were used in more than half of the apps—1.1 Goal setting (behavior; 52% of apps), 2.2 Feedback on behavior (54%), 2.3 Self-monitoring of behavior (72%), and 4.1 Instructions on how to perform the behavior (54%). The mean (and median) number of BCTs per app was 5 (range: 0-11), and the most common number of BCTs per app was 6.

An exploratory assessment of the effectiveness of each BCT was conducted by associating the use of BCTs in each app with the effectiveness of that app, so that a count of how many times the BCT was associated with significant evidence versus no evidence could be made. There was at best mixed evidence for all of the BCTs used. Only 4 BCTs (1.6 Discrepancy between current behavior and goal, 4.2 Information about antecedents, 6.1 Demonstration of the behavior, and 12.5 Adding objects to the environment) had more significant evidence than not, but only by 1 study. These 4 BCTs were also only used in at most 2 apps. The most frequently used BCTs were all associated two to three times more with studies that found no significant effect compared with those that found a significant effect on the specified outcomes.

Half of the studies (26/52) mentioned the specific behavioral theories that were considered when developing the app, and there was a lot of variety. A total of 23 different theories were referenced, with social cognitive theory and behavior change theory being referenced most frequently (8 and 5 times, respectively), with the remaining 21 theories having no more than two mentions each. Only 5 of these 26 studies found significant evidence in favor of the app. Of these 5, 2 used social cognitive theory, 2 used bahavior change theory, and 1 used the capability, opportunity, motivation, behavior framework and the behavior change wheel.

Engagement and Adherence

Engagement and adherence outcomes were reported by 39 of the studies. Of these, 18 reported completion or retention rates, and 26 reported the app use data. The mean completion rate across studies was 83.3% (range: 45%-97.1%), with 8 studies reporting a completion rate above 90%. There was significant variability in what app use measures were reported and how they were used to evaluate adherence. A total of 4 studies reported that the app group was significantly more engaged with their intervention than the control group, and 2 more reported high use in the app group but not whether the difference was significant. A total of 9 studies reported a usage percentage greater than 60%.

Feasibility and Usability

A total of 15 studies reported on usability (n=13), feasibility (n=1), or both (n=1). A total of 3 of these studies reported that the intervention app was rated significantly better than the control, and 9 more reported high ease of use (>70% of participants rated highly). There does not appear to be any relationship between usability and effectiveness, given the generally high usability ratings and overall low effectiveness. However, as less than a third of studies reported usability, this analysis should be treated with caution.

Participant Satisfaction

A total of 21 studies reported participant satisfaction. Of these studies, 2 reported significantly higher ratings for the intervention app than the control, and 6 more reported that more than 70% of the participants rated the app as satisfying, helpful, or enjoyable. Only 1 app (Crush the Crave) was found to have significantly lower helpfulness and satisfaction ratings than the control. The rest had mixed feedback or no significant differences between the ratings for the app and control.

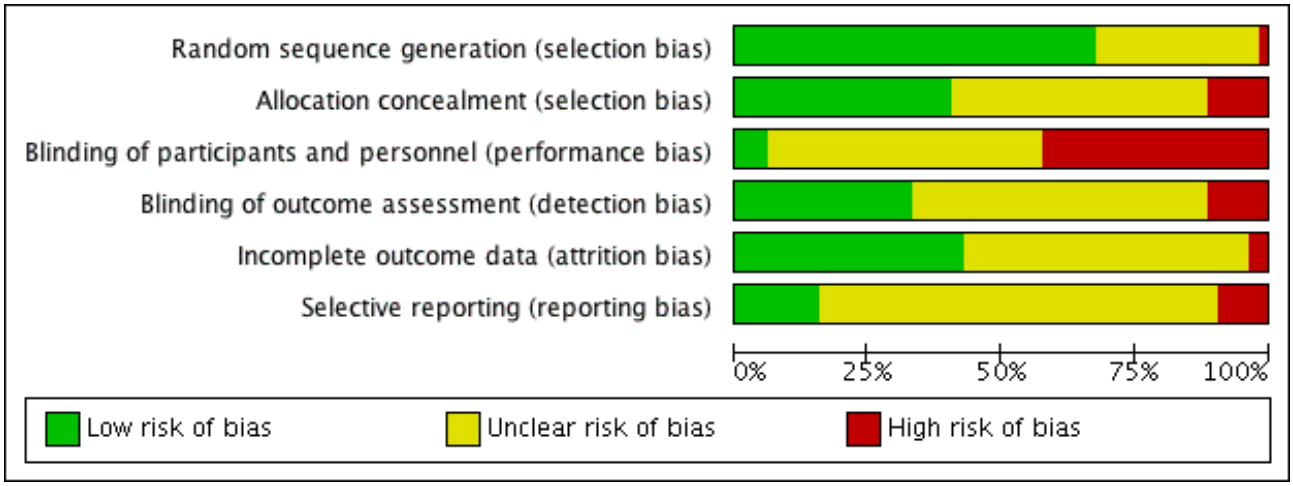
Risk of Bias Assessment

The evaluation of risk of bias for all 52 studies was conducted using the Cochrane Collaboration Risk of Bias tool [20], and the results were summarized using the RevMan 5.3 software (Figures 2 and 3) [73]. Two-thirds of the studies (35/52) properly reported random sequence generation [36].

About 40% (21/52) of the studies reported satisfactory allocation concealment and either very low attrition or no significant differences in attrition between groups (22/52), meaning that the risk of incomplete outcome data was low. About a third of studies (17/52) reported blinding of outcome assessors, but only 3 studies reported blinding of patients and personnel. This was predominantly because the nature of mobile app interventions made blinding of participants difficult. A total of 42% (22/52) of the studies had a high risk for the blinding of participants and personnel, predominantly because they reported that blinding was not possible. However, only 15% (8/52) of the studies could be established as having a low risk of selective outcome, mostly because a preregistration or study protocol could not be found to compare reported outcomes with.

Figure 2. Risk of bias summary: the review authors' judgements about each risk of bias item for each included study.

Figure 3. Risk of bias graph: the review authors' judgements about each risk of bias item presented as percentages across all included studies.



Discussion

Principal Findings

Developments in technology have made it easier for patients to play an active role in their health, and there are thousands of mobile apps designed to help people improve their health behaviors [4]. However, their effectiveness in changing health behaviors and outcomes has not been convincingly established. In this systematic review, we examined 52 randomized controlled trials evaluating the effectiveness of mobile health apps. Overall, there was little significant evidence supporting the effectiveness of mobile health apps for any of those outcomes. This was especially true for patient health outcomes; out of all the studies that examined them, there was only 1 app that was significantly better than the control. There was slightly more evidence for the effectiveness of apps in changing the health behavior outcomes—apps performed significantly better than controls on a quarter of the measured outcomes—but a majority of outcomes reported no significant differences between groups.

This is a different finding than previous systematic reviews. Both Han and Lee [12] and Payne et al [15] reported that the majority of apps reviewed were effective at improving health-related behaviors. It is possible that this difference is because of the way study results were interpreted. Only considering studies that found a significant difference between groups in favor of the app as significant evidence of effectiveness is a conservative interpretation. However, randomized controlled trials are the gold standard for evaluating interventions because randomization allows differences in outcomes between groups to be attributed to the intervention [74]. To conclusively support the claim that mobile health apps are a useful tool for changing behavior and health outcomes, users should show greater improvement with the app than a comparator or control. The majority of the apps evaluated in this review were not significantly more effective in achieving their purposes than controls or alternative interventions.

When significant and some evidence categories are both taken into account, there was moderately strong evidence supporting the effectiveness of apps in changing certain target behaviors, notably, healthy food choices (6/8, 75%), step count (5/7, 71%), and reducing sedentary behavior (3/3, 100%). This suggests that mobile apps do have the potential to improve some health behaviors. However, the number of studies that examined each of these outcomes was small. In addition, this effect was not reflected in any of the participant health outcomes. Possibly the intervention durations were too short for any meaningful clinical change, although half of the studies lasted longer than 3 months, and modest weight loss (5%-10%) can be observed within 3 to 6 months [75]. Even when the behavior change was significantly greater in the app than the control group, it may not have been enough of a change to induce an observable effect in any of the health outcomes measured over time.

Before delving too deeply into clinical outcomes, however, it is first essential to have a measurable behavioral effect. Identifying the BCTs that are most effective in promoting and maintaining positive health behavior change is crucial for the development of mobile apps that will significantly improve health behaviors and outcomes [76]. However, determining which BCTs, and combinations of BCTs, are most effective in specific contexts is a complex process, and a valid method of determining the degree of confidence of BCT effectiveness is yet to be established [76]. To make this even more difficult, most of the studies did not report the BCTs used, and they had to be inferred from the descriptions of the apps' features. Self-monitoring of behavior was the most commonly used BCT (72% of apps included a self-monitoring function). Behavioral goal setting, feedback on behavior, and instructions on how to perform the behavior were also included in more than half of the apps. This is consistent with the findings of Payne et al's [15] systematic review of mobile health apps, which found that most studies included goal setting, self-monitoring, and social support constructs. Social support was not as prominent in the apps studied in this review, with only 28% having an identifiable social support feature. If there is a disconnect between the BCTs that are most effective and those that are most frequently used, it could explain the lack of behavioral change.

For all 4 of the most frequently used BCTs, there were far more studies with no evidence than significant evidence, with less than a fifth of the studies of apps using those individual BCTs finding a significant effect (range: 15%-19% per BCT). This is similar to the low overall amount of significant evidence

supporting the apps' effectiveness, which is not surprising. Self-monitoring has been positively associated with behavior change in the literature, though the results are heterogenous [77,78]. Therefore, why are these studies not finding much support for their effectiveness in changing behavior?

Intuitively, a greater number of BCTs might seem more likely to improve health behaviors, or at least provide a wider range of motivating options so that users can choose the ones that work best for them. However, there was a similar average number of BCTs per app for each of the evidence groups (no, some, and significant evidence). The no evidence group actually had a higher average number of BCTs per app (5.9 vs 4.4 in the other two groups) but two-tailed *t* tests showed that none of the differences between the groups were significant ($P>.05$). This lack of an association between the number of BCTs and effectiveness is consistent with previous studies [79].

It is possible that the apps were not appropriately implementing the BCTs in their features. The lack of theoretical bases for half of the apps suggests that BCTs were not considered in many of them. However, only 4 of the 26 studies that did reference behavioral theories (actually 5 out of 27 studies, but 2 studies reported the same experiment [22,47]) found significant evidence for the app they studied, which is no better than the overall group of studies. This suggests that the inclusion of theory is not sufficient to find an effective result. The variety of theories used could also be a factor; further research should determine which theoretical bases are most strongly linked to behavior change. A more in-depth evaluation of the use and effectiveness of BCTs—as individual factors and in combination—in mobile health apps is necessary to understand why the majority of studies are not finding significant behavioral or health advantages for mobile app interventions. This is especially important because of the ubiquity of health apps that are available and being used and the urgent public health need to improve health behaviors to address increasing health care costs and improve healthy aging [80].

Quality of the Evidence

After analyzing the risk of bias of the included studies, only the category of performance bias (blinding of participants and personnel) had a generally high risk of bias (22/52, 42%). Overall, however, just over half of the potential areas for bias had an unclear risk. Therefore, to improve the quality of studies and to make bias assessments more clear and useful, researchers should improve reporting of their methods, so that the risk of bias can be assessed more accurately. Out of all 52 studies, only 3 had 5 or more areas of bias categorized as low risk [27]. High-quality studies are needed to make a valid evaluation of the effectiveness of apps, so that there is less risk of poor study methodologies confusing the conclusions.

Limitations

One limitation of this review is that a meta-analysis could not be conducted because of the heterogeneity of the studies and their reported outcomes. However, a proper meta-analysis would make the effectiveness of mobile health apps easier to determine and quantify. Another limitation is that the focus was limited to just five health behaviors. There are many mobile health apps designed to help patients manage chronic conditions such as diabetes, depression, and asthma. Therefore, the results of this review cannot be generalized to all health behaviors. In addition, this review only considered published randomized controlled trials. This may have missed more recently developed apps that have not progressed to that stage of testing yet, or that might not have been tested in an academic context. It may also overrepresent studies where an effect was found, as the grey literature was not searched for studies that may have found null results and not been published.

Future Directions

An important future direction for research—and app development—is to examine more closely the theoretical basis of mobile health apps, which BCTs they are using, and how those BCTs are implemented. This is a crucial element in determining why mobile health apps are not consistently succeeding in improving health behaviors. If the most effective BCTs, and combinations thereof, can be identified, mobile health apps have the potential to advance preventive health care globally. Once consistent and effective means of motivating behavior change have been identified, the relationship between health behaviors and health outcomes should be reassessed and, if necessary, improved. To complement this, it is important for researchers to improve their reporting, so that the risk of bias of studies can be accurately assessed and only high-quality studies can be included in analyses.

Conclusions

The purpose of this systematic review was to examine the effectiveness of mobile apps to improve health behaviors and outcomes and the inclusion and effectiveness of BCTs. Although apps generally had relatively high engagement, usability ratings, and user acceptability and satisfaction, the significance of evidence for delivering behavior change outcomes assessed was nominal. This study built on previous systematic reviews to provide an updated and comprehensive examination of current mobile health apps for the general population. It extended the literature by examining the relationship between BCTs and app effectiveness. In addition, this systematic review evaluated effectiveness more stringently than previous reviews to provide a balanced perspective on current app effectiveness and identify areas for improvement. Further research is needed to identify the behavior change theories and specific BCTs best suited to promote and maintain positive health behavior change through mobile app interventions. A reliable method of analyzing BCT effectiveness and more experiments comparing how behavior change outcomes differ depending on the combinations of BCTs used would be a useful next step. Given the inconsistent results of studies of mobile health app effectiveness, a greater integration of theory into app development and comparative examination of theories and BCTs in those apps will help drive innovation and the creation of more effective mobile health apps.

Acknowledgments

The authors would like to thank the outreach librarian Liz Callow for her assistance in developing search terms and reviewing search strategy. MM, CL, CDC, MHVVV, and EM are supported by grants from EIT Health (Grant 18654) and Promoting the Internet of Things via Collaborations between HEIS and Industry (PITCH-IN) grant.

Authors' Contributions

EM and MM conceived the study topic and designed the review protocol. MM and CL screened the studies. MM conducted the data extraction and risk of bias assessment, which were validated by CC. The systematic review was written by MM with revisions from MV and EM.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist.
[DOC File , 64 KB - mhealth_v8i3e17046_app1.doc]

Multimedia Appendix 2
Search Terms.
[DOCX File , 15 KB - mhealth_v8i3e17046_app2.docx]

Multimedia Appendix 3
Search queries and the number of results for each database.
[DOCX File , 15 KB - mhealth_v8i3e17046_app3.docx]

Multimedia Appendix 4
Summary of study characteristics.
[DOCX File , 23 KB - mhealth_v8i3e17046_app4.docx]

Multimedia Appendix 5
Behaviour Change Techniques used in the apps studied.
[ZIP File (Zip Archive), 3 KB - mhealth_v8i3e17046_app5.zip]

References

1. Clancy CM. Patient engagement in health care. Health Serv Res 2011 Apr;46(2):389-393 [FREE Full text] [doi: 10.1111/j.1475-6773.2011.01254.x] [Medline: 21371026]
2. Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF. The impact of information technology on patient engagement and health behavior change: a systematic review of the literature. JMIR Med Inform 2016 Jan 21;4(1):e1 [FREE Full text] [doi: 10.2196/medinform.4514] [Medline: 26795082]
3. Stephens TN, Joerin A, Rauws M, Werk LN. Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot. Transl Behav Med 2019 May 16;9(3):440-447. [doi: 10.1093/tbm/ibz043] [Medline: 31094445]
4. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using health and well-being apps for behavior change: a systematic search and rating of apps. JMIR Mhealth Uhealth 2019 Jul 4;7(7):e11926 [FREE Full text] [doi: 10.2196/11926] [Medline: 31274112]
5. Jiang H, Li M, Wen LM, Baur L, He G, Ma X, et al. A community-based short message service intervention to improve mothers' feeding practices for obesity prevention: quasi-experimental study. JMIR Mhealth Uhealth 2019 Jun 3;7(6):e13828 [FREE Full text] [doi: 10.2196/13828] [Medline: 31162133]
6. Ferrara G, Kim J, Lin S, Hua J, Seto E. A focused review of smartphone diet-tracking apps: usability, functionality, coherence with behavior change theory, and comparative validity of nutrient intake and energy estimates. JMIR Mhealth Uhealth 2019 May 17;7(5):e9232 [FREE Full text] [doi: 10.2196/mhealth.9232] [Medline: 31102369]
7. Trifan A, Oliveira M, Oliveira JL. Passive sensing of health outcomes through smartphones: systematic review of current solutions and possible limitations. JMIR Mhealth Uhealth 2019 Aug 23;7(8):e12649 [FREE Full text] [doi: 10.2196/12649] [Medline: 31444874]

8. Harari GM, Lane ND, Wang R, Crosier BS, Campbell AT, Gosling SD. Using smartphones to collect behavioral data in psychological science: opportunities, practical considerations, and challenges. Perspect Psychol Sci 2016 Nov;11(6):838-854 [FREE Full text] [doi: 10.1177/1745691616650285] [Medline: 27899727]
9. Cockerham WC, Dingwall R, Quah SR. Health behavior. In: Cockerham WC, Dingwall R, Quah SR, editors. The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society. Chichester, UK: John Wiley & Sons Ltd; 2014.
10. Conner M, Norman P. Health behaviour: current issues and challenges. Psychol Health 2017 Aug;32(8):895-906. [doi: 10.1080/08870446.2017.1336240] [Medline: 28612656]
11. Pinder C, Vermeulen J, Cowan BR, Beale R. Digital behaviour change interventions to break and form habits. ACM Trans Comput-Hum Interact 2018;25(3):1-66 [FREE Full text] [doi: 10.1145/3196830]
12. Han M, Lee E. Effectiveness of mobile health application use to improve health behavior changes: a systematic review of randomized controlled trials. Healthc Inform Res 2018 Jul;24(3):207-226 [FREE Full text] [doi: 10.4258/hir.2018.24.3.207] [Medline: 30109154]
13. 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]
14. Hagger MS, Weed M. DEBATE: Do interventions based on behavioral theory work in the real world? Int J Behav Nutr Phys Act 2019 Apr 25;16(1):36 [FREE Full text] [doi: 10.1186/s12966-019-0795-4] [Medline: 31023328]
15. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
16. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. Br Med J 2015 Jan 2;349:g7647 [FREE Full text] [doi: 10.1136/bmj.g7647] [Medline: 25555855]
17. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3 [FREE Full text] [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
18. 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]
19. Hedin B, Katzeff C, Eriksson E, Pargman D. A systematic review of digital behaviour change interventions for more sustainable food consumption. Sustainability 2019;11(9):2638 [FREE Full text] [doi: 10.3390/su11092638]
20. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Br Med J 2011 Oct 18;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
21. López D, Torres M, Vélez J, Grullon J, Negrón E, Pérez CM, et al. Development and evaluation of a nutritional smartphone application for making smart and healthy choices in grocery shopping. Healthc Inform Res 2017 Jan;23(1):16-24 [FREE Full text] [doi: 10.4258/hir.2017.23.1.16] [Medline: 28261527]
22. Mummah S, Robinson TN, Mathur M, Farzinkhou S, Sutton S, Gardner CD. Effect of a mobile app intervention on vegetable consumption in overweight adults: a randomized controlled trial. Int J Behav Nutr Phys Act 2017 Sep 15;14(1):125 [FREE Full text] [doi: 10.1186/s12966-017-0563-2] [Medline: 28915825]
23. Clarke P, Evans SH, Neffa-Creech D. Mobile app increases vegetable-based preparations by low-income household cooks: a randomized controlled trial. Public Health Nutr 2019 Mar;22(4):714-725 [FREE Full text] [doi: 10.1017/S1368980018003117] [Medline: 30472970]
24. Walsh JC, Corbett T, Hogan M, Duggan J, McNamara A. An mHealth intervention using a smartphone app to increase walking behavior in young adults: a pilot study. JMIR Mhealth Uhealth 2016 Sep 22;4(3):e109 [FREE Full text] [doi: 10.2196/mhealth.5227] [Medline: 27658677]
25. Zhou M, Fukuoka Y, Mintz Y, Goldberg K, Kaminsky P, Flowers E, et al. Evaluating machine learning-based automated personalized daily step goals delivered through a mobile phone app: randomized controlled trial. JMIR Mhealth Uhealth 2018 Jan 25;6(1):e28 [FREE Full text] [doi: 10.2196/mhealth.9117] [Medline: 29371177]
26. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app 'Calm' to reduce stress among college students: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]
27. Crane D, Garnett C, Michie S, West R, Brown J. A smartphone app to reduce excessive alcohol consumption: Identifying the effectiveness of intervention components in a factorial randomised control trial. Sci Rep 2018 Mar 12;8(1):4384 [FREE Full text] [doi: 10.1038/s41598-018-22420-8] [Medline: 29531280]
28. BinDhim NF, McGeechan K, Trevena L. Smartphone Smoking Cessation Application (SSC App) trial: a multicountry double-blind automated randomised controlled trial of a smoking cessation decision-aid 'app'. BMJ Open 2018 Jan 21;8(1):e017105 [FREE Full text] [doi: 10.1136/bmjopen-2017-017105] [Medline: 29358418]
29. Arrogi A, Bogaerts A, Seghers J, Devloo K, Abeele VV, Geurts L, et al. Evaluation of stAPP: a smartphone-based intervention to reduce prolonged sitting among Belgian adults. Health Promot Int 2019 Feb 1;34(1):16-27. [doi: 10.1093/heapro/dax046] [Medline: 28973149]

30. Cowdery J, Majeske P, Frank R, Brown D. Exergame apps and physical activity: the results of the ZOMBIE trial. Am J Health Educ 2015;46(4):216-222. [doi: 10.1080/19325037.2015.1043063]
31. Direito A, Jiang Y, Whittaker R, Maddison R. Apps for IMproving FITness and increasing physical activity among young people: the AIMFIT pragmatic randomized controlled trial. J Med Internet Res 2015 Aug 27;17(8):e210 [FREE Full text] [doi: 10.2196/jmir.4568] [Medline: 26316499]
32. Finkelstein J, Bedra M, Li X, Wood J, Ouyang P. Mobile app to reduce inactivity in sedentary overweight women. Stud Health Technol Inform 2015;216:89-92. [Medline: 26262016]
33. Fukuoka Y, Haskell W, Lin F, Vittinghoff E. Short- and long-term effects of a mobile phone app in conjunction with brief in-person counseling on physical activity among physically inactive women: the mPED randomized clinical trial. JAMA Netw Open 2019 May 3;2(5):e194281 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.4281] [Medline: 31125101]
34. Gabbiadini A, Greitemeyer T. Fitness mobile apps positively affect attitudes, perceived behavioral control and physical activities. J Sports Med Phys Fitness 2019 Mar;59(3):407-414. [doi: 10.23736/S0022-4707.18.08260-9] [Medline: 29619794]
35. Glynn LG, Hayes PS, Casey M, Glynn F, Alvarez-Iglesias A, Newell J, et al. Effectiveness of a smartphone application to promote physical activity in primary care: the SMART MOVE randomised controlled trial. Br J Gen Pract 2014 Jul;64(624):e384-e391 [FREE Full text] [doi: 10.3399/bjgp14X680461] [Medline: 24982490]
36. Harries T, Eslambolchilar P, Rettie R, Stride C, Walton S, van Woerden HC. Effectiveness of a smartphone app in increasing physical activity amongst male adults: a randomised controlled trial. BMC Public Health 2016 Sep 2;16:925 [FREE Full text] [doi: 10.1186/s12889-016-3593-9] [Medline: 27590255]
37. King AC, Hekler EB, Grieco LA, Winter SJ, Sheats JL, Buman MP, et al. Effects of three motivationally targeted mobile device applications on initial physical activity and sedentary behavior change in midlife and older adults: a randomized trial. PLoS One 2016;11(6):e0156370 [FREE Full text] [doi: 10.1371/journal.pone.0156370] [Medline: 27352250]
38. Simons D, de Bourdeaudhuij I, Clarys P, de Cocker K, Vandelanotte C, Deforche B. Effect and process evaluation of a smartphone app to promote an active lifestyle in lower educated working young adults: cluster randomized controlled trial. JMIR Mhealth Uhealth 2018 Aug 24;6(8):e10003 [FREE Full text] [doi: 10.2196/10003] [Medline: 30143477]
39. Voth EC, Oelke ND, Jung ME. A theory-based exercise app to enhance exercise adherence: a pilot study. JMIR Mhealth Uhealth 2016 Jun 15;4(2):e62 [FREE Full text] [doi: 10.2196/mhealth.4997] [Medline: 27307134]
40. Blackburne T, Rodriguez A, Johnstone SJ. A serious game to increase healthy food consumption in overweight or obese adults: randomized controlled trial. JMIR Serious Games 2016 Jul 13;4(2):e10 [FREE Full text] [doi: 10.2196/games.5708] [Medline: 27417192]
41. Brindal E, Hendrie GA, Freyne J, Noakes M. Incorporating a static versus supportive mobile phone app into a partial meal replacement program with face-to-face support: randomized controlled trial. JMIR Mhealth Uhealth 2018 Apr 18;6(4):e41 [FREE Full text] [doi: 10.2196/mhealth.7796] [Medline: 29669704]
42. Dunn CG, Turner-McGrievy GM, Wilcox S, Hutto B. Dietary self-monitoring through calorie tracking but not through a digital photography app is associated with significant weight loss: the 2SMART pilot study-a 6-month randomized trial. J Acad Nutr Diet 2019 Sep;119(9):1525-1532. [doi: 10.1016/j.jand.2019.03.013] [Medline: 31155474]
43. Ipjian ML, Johnston CS. Smartphone technology facilitates dietary change in healthy adults. Nutrition 2017 Jan;33:343-347. [doi: 10.1016/j.nut.2016.08.003] [Medline: 27742102]
44. Kakoschke N, Hawker C, Castine B, de Courten B, Verdejo-Garcia A. Smartphone-based cognitive bias modification training improves healthy food choice in obesity: a pilot study. Eur Eat Disord Rev 2018 Sep;26(5):526-532. [doi: 10.1002/erv.2622] [Medline: 30003634]
45. Kliemann N, Croker H, Johnson F, Beeken RJ. Development of the top tips habit-based weight loss app and preliminary indications of its usage, effectiveness, and acceptability: mixed-methods pilot study. JMIR Mhealth Uhealth 2019 May 10;7(5):e12326 [FREE Full text] [doi: 10.2196/12326] [Medline: 31094352]
46. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-12 [FREE Full text] [doi: 10.7326/M13-3005] [Medline: 25402403]
47. Mummah SA, Mathur M, King AC, Gardner CD, Sutton S. Mobile technology for vegetable consumption: a randomized controlled pilot study in overweight adults. JMIR Mhealth Uhealth 2016 May 18;4(2):e51 [FREE Full text] [doi: 10.2196/mhealth.5146] [Medline: 27193036]
48. Palacios C, Torres M, López D, Trak-Fellermeier MA, Coccia C, Pérez CM. Effectiveness of the nutritional app 'MyNutriCart' on food choices related to purchase and dietary behavior: a pilot randomized controlled trial. Nutrients 2018 Dec 12;10(12):pii: E1967 [FREE Full text] [doi: 10.3390/nu10121967] [Medline: 30545125]
49. Whitelock V, Kersbergen I, Higgs S, Aveyard P, Halford JC, Robinson E. A smartphone based attentive eating intervention for energy intake and weight loss: results from a randomised controlled trial. BMC Public Health 2019 May 21;19(1):611 [FREE Full text] [doi: 10.1186/s12889-019-6923-x] [Medline: 31113400]
50. Brindal E, Hendrie GA, Freyne J, Noakes M. A mobile phone app designed to support weight loss maintenance and well-being (MotiMate): randomized controlled trial. JMIR Mhealth Uhealth 2019 Sep 4;7(9):e12882 [FREE Full text] [doi: 10.2196/12882] [Medline: 31486407]

51. Eyles H, McLean R, Neal B, Jiang Y, Doughty RN, McLean R, et al. A salt-reduction smartphone app supports lower-salt food purchases for people with cardiovascular disease: Findings from the SaltSwitch randomised controlled trial. Eur J Prev Cardiol 2017 Sep;24(13):1435-1444. [doi: 10.1177/2047487317715713] [Medline: 28631933]
52. Garcia-Ortiz L, Recio-Rodriguez JI, Agudo-Conde C, Patino-Alonso MC, Maderuelo-Fernandez J, Gento IR, EVIDENT Investigators Group, Mobilizing Minds Research Group. Long-term effectiveness of a smartphone app for improving healthy lifestyles in general population in primary care: randomized controlled trial (Evident II Study). JMIR Mhealth Uhealth 2018 Apr 27;6(4):e107 [FREE Full text] [doi: 10.2196/mhealth.9218] [Medline: 29702473]
53. Gomez-Marcos MA, Patino-Alonso MC, Recio-Rodriguez JI, Agudo-Conde C, Romaguera-Bosch M, Magdalena-Gonzalez O, on behalf the EVIDENT Investigators 11. Short- and long-term effectiveness of a smartphone application for improving measures of adiposity: A randomised clinical trial - EVIDENT II study. Eur J Cardiovasc Nurs 2018 Aug;17(6):552-562. [doi: 10.1177/1474515118761870] [Medline: 29488798]
54. Gonzalez-Sanchez J, Recio-Rodriguez JI, Fernandez-delRio A, Sanchez-Perez A, Magdalena-Belio JF, Gomez-Marcos MA, On behalf the EVIDENT Investigators group. Using a smartphone app in changing cardiovascular risk factors: a randomized controlled trial (EVIDENT II study). Int J Med Inform 2019 May;125:13-21. [doi: 10.1016/j.ijmedinf.2019.02.004] [Medline: 30914176]
55. Hurkmans E, Matthys C, Bogaerts A, Scheys L, Devloo K, Seghers J. Face-to-face versus mobile versus blended weight loss program: randomized clinical trial. JMIR Mhealth Uhealth 2018 Jan 11;6(1):e14 [FREE Full text] [doi: 10.2196/mhealth.7713] [Medline: 29326093]
56. Lin P, Grambow S, Intille S, Gallis JA, Lazenka T, Bosworth H, et al. The association between engagement and weight loss through personal coaching and cell phone interventions in young adults: randomized controlled trial. JMIR Mhealth Uhealth 2018 Oct 18;6(10):e10471 [FREE Full text] [doi: 10.2196/10471] [Medline: 30341051]
57. 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 May 14;3(2):e42 [FREE Full text] [doi: 10.2196/mhealth.4160] [Medline: 25977197]
58. Recio-Rodriguez JI, Agudo-Conde C, Martin-Cantera C, González-Viejo MN, Fernandez-Alonso MD, Arietaleanizbeaskoa MS, EVIDENT Investigators. Short-term effectiveness of a mobile phone app for increasing physical activity and adherence to the mediterranean diet in primary care: a randomized controlled trial (Evident II Study). J Med Internet Res 2016 Dec 19;18(12):e331 [FREE Full text] [doi: 10.2196/jmir.6814] [Medline: 27993759]
59. Svetkey LP, Batch BC, Lin P, Intille SS, Corsino L, Tyson CC, et al. Cell phone intervention for you (CITY): a randomized, controlled trial of behavioral weight loss intervention for young adults using mobile technology. Obesity (Silver Spring) 2015 Nov;23(11):2133-2141 [FREE Full text] [doi: 10.1002/oby.21226] [Medline: 26530929]
60. van Beurden SB, Smith JR, Lawrence NS, Abraham C, Greaves CJ. Feasibility randomized controlled trial of ImpulsePal: smartphone app-based weight management intervention to reduce impulsive eating in overweight adults. JMIR Form Res 2019 Apr 30;3(2):e11586 [FREE Full text] [doi: 10.2196/11586] [Medline: 31038464]
61. Bakker D, Kazantzis N, Rickwood D, Rickard N. A randomized controlled trial of three smartphone apps for enhancing public mental health. Behav Res Ther 2018 Oct;109:75-83. [doi: 10.1016/j.brat.2018.08.003] [Medline: 30125790]
62. Champion L, Economides M, Chandler C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: a pilot randomised controlled trial. PLoS One 2018;13(12):e0209482 [FREE Full text] [doi: 10.1371/journal.pone.0209482] [Medline: 30596696]
63. Hides L, Dingle G, Quinn C, Stoyanov SR, Zelenko O, Tjondronegoro D, et al. Efficacy and outcomes of a music-based emotion regulation mobile app in distressed young people: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11482 [FREE Full text] [doi: 10.2196/11482] [Medline: 30664457]
64. Baskerville NB, Struik LL, Guindon GE, Norman CD, Whittaker R, Burns C, et al. Effect of a mobile phone intervention on quitting smoking in a young adult population of smokers: randomized controlled trial. JMIR Mhealth Uhealth 2018 Oct 23;6(10):e10893 [FREE Full text] [doi: 10.2196/10893] [Medline: 30355563]
65. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 1;143:87-94 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]
66. Buller DB, Borland R, Bettinghaus EP, Shane JH, Zimmerman DE. Randomized trial of a smartphone mobile application compared to text messaging to support smoking cessation. Telemed J E Health 2014 Mar;20(3):206-214 [FREE Full text] [doi: 10.1089/tmj.2013.0169] [Medline: 24350804]
67. Crane D, Ubhi HK, Brown J, West R. Relative effectiveness of a full versus reduced version of the 'Smoke Free' mobile application for smoking cessation: an exploratory randomised controlled trial. F1000Res 2018;7:1524 [FREE Full text] [doi: 10.12688/f1000research.16148.2] [Medline: 30728950]
68. 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 May 26;5(5):e74 [FREE Full text] [doi: 10.2196/mhealth.6252] [Medline: 28550004]

69. Krishnan N, Elf JL, Chon S, Golub JE. COach2Quit: a pilot randomized controlled trial of a personal carbon monoxide monitor for smoking cessation. Nicotine Tob Res 2019 Oct 26;21(11):1573-1577 [FREE Full text] [doi: 10.1093/ntr/nty182] [Medline: 30169740]
70. McClure JB, Anderson ML, Bradley K, An LC, Catz SL. Evaluating an adaptive and interactive mHealth smoking cessation and medication adherence program: a randomized pilot feasibility study. JMIR Mhealth Uhealth 2016 Aug 3;4(3):e94 [FREE Full text] [doi: 10.2196/mhealth.6002] [Medline: 27489247]
71. Peiris D, Wright L, News M, Rogers K, Redfern J, Chow C, et al. A smartphone app to assist smoking cessation among aboriginal Australians: findings from a pilot randomized controlled trial. JMIR Mhealth Uhealth 2019 Apr 2;7(4):e12745 [FREE Full text] [doi: 10.2196/12745] [Medline: 30938691]
72. Bertholet N, Godinho A, Cunningham JA. Smartphone application for unhealthy alcohol use: pilot randomized controlled trial in the general population. Drug Alcohol Depend 2019 Feb 1;195:101-105. [doi: 10.1016/j.drugalcdep.2018.12.002] [Medline: 30611977]
73. The Cochrane Collaboration. Copenhagen: The Nordic Cochrane Centre; 2014. Review Manager (RevMan) URL: https://community.cochrane.org/help/tools-and-software/revman-5 [accessed 2020-02-26]
74. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. Br J Obstet Gynaecol 2018 Dec;125(13):1716 [FREE Full text] [doi: 10.1111/1471-0528.15199] [Medline: 29916205]
75. Nymo S, Coutinho SR, Torgersen LH, Bomo OJ, Haugvaldstad I, Truby H, et al. Timeline of changes in adaptive physiological responses, at the level of energy expenditure, with progressive weight loss. Br J Nutr 2018 Jul;120(2):141-149 [FREE Full text] [doi: 10.1017/S0007114518000922] [Medline: 29733003]
76. Michie S, West R, Sheals K, Godinho CA. Evaluating the effectiveness of behavior change techniques in health-related behavior: a scoping review of methods used. Transl Behav Med 2018 Mar 1;8(2):212-224 [FREE Full text] [doi: 10.1093/tbm/ibx019] [Medline: 29381786]
77. Teasdale N, Elhussein A, Butcher F, Piernas C, Cowburn G, Hartmann-Boyce J, et al. Systematic review and meta-analysis of remotely delivered interventions using self-monitoring or tailored feedback to change dietary behavior. Am J Clin Nutr 2018 Feb 1;107(2):247-256 [FREE Full text] [doi: 10.1093/ajcn/nqx048] [Medline: 29529158]
78. Chambliss HO, Huber RC, Finley CE, McDoniel SO, Kitzman-Ulrich H, Wilkinson WJ. Computerized self-monitoring and technology-assisted feedback for weight loss with and without an enhanced behavioral component. Patient Educ Couns 2011 Dec;85(3):375-382. [doi: 10.1016/j.pec.2010.12.024] [Medline: 21295433]
79. Dombrowski SU, Sniehotta FF, Avenell A, Johnston M, MacLennan G, Araújo-Soares V. Identifying active ingredients in complex behavioural interventions for obese adults with obesity-related co-morbidities or additional risk factors for co-morbidities: a systematic review. Health Psychol Rev 2012;6(1):7-32. [doi: 10.1080/17437199.2010.513298]
80. Wendimagegn NF, Bezuidenhout MC. Integrating promotive, preventive, and curative health care services at hospitals and health centers in Addis Ababa, Ethiopia. J Multidiscip Healthc 2019;12:243-255 [FREE Full text] [doi: 10.2147/JMDH.S193370] [Medline: 31040687]

Abbreviations

BCT: behavior change technique

Edited by G Eysenbach; submitted 13.11.19; peer-reviewed by S Sawesi, I Mircheva; comments to author 29.11.19; revised version received 03.12.19; accepted 26.01.20; published 18.03.20.

Please cite as:
Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E
Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review
JMIR Mhealth Uhealth 2020;8(3):e17046
URL: http://mhealth.jmir.org/2020/3/e17046/
doi:10.2196/17046
PMID:32186518

Review

The Effectiveness of Self-Management of Hypertension in Adults Using Mobile Health: Systematic Review and Meta-Analysis

Ran Li[1], MSc; Ning Liang[2], MSc, PhD; Fanlong Bu[3], MD, PhD; Therese Hesketh[1,4], PhD

[1]Center of Global Health, School of Public Health, School of Medicine, Zhejiang University, Hangzhou, China

[2]Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China

[3]Center for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China

[4]Institute of Global Health, University College London, London, United Kingdom

Corresponding Author:
Therese Hesketh, PhD
Institute of Global Health
University College London
30 Guilford St
London, WC1N 1EH
United Kingdom
Phone: 44 0207 905 2253
Email: t.hesketh@ucl.ac.uk

Abstract

Background: Effective treatment of hypertension requires careful self-management. With the ongoing development of mobile technologies and the scarcity of health care resources, mobile health (mHealth)–based self-management has become a useful treatment for hypertension, and its effectiveness has been assessed in many trials. However, there is a paucity of comprehensive summaries of the studies using both qualitative and quantitative methods.

Objective: This systematic review aimed to measure the effectiveness of mHealth in improving the self-management of hypertension for adults. The outcome measures were blood pressure (BP), BP control, medication adherence, self-management behavior, and costs.

Methods: A systematic search was conducted using 5 electronic databases. The snowballing method was used to scan the reference lists of relevant studies. Only peer-reviewed randomized controlled trials (RCTs) published between January 2010 and September 2019 were included. Data extraction and quality assessment were performed by 3 researchers independently, adhering to the validation guideline and checklist. Both a meta-analysis and a narrative synthesis were carried out.

Results: A total of 24 studies with 8933 participants were included. Of these, 23 studies reported the clinical outcome of BP, 12 of these provided systolic blood pressure (SBP) and diastolic blood pressure (DBP) data, and 16 articles focused on change in self-management behavior and medication adherence. All 24 studies were included in the narrative synthesis. According to the meta-analysis, a greater reduction in both SBP and DBP was observed in the mHealth intervention groups compared with control groups, −3.78 mm Hg (P<.001; 95% CI −4.67 to −2.89) and −1.57 mm Hg (P<.001; 95% CI −2.28 to −0.86), respectively. Subgroup analyses showed consistent reductions in SBP and DBP across different frequencies of reminders, interactive patterns, intervention functions, and study duration subgroups. A total of 16 studies reported better medication adherence and behavioral change in the intervention groups, while 8 showed no significant change. Six studies included an economic evaluation, which drew inconsistent conclusions. However, potentially long-term financial benefits were mentioned in all economic evaluations. All studies were assessed to be at high risk of bias.

Conclusions: This review found that mHealth self-management interventions were effective in BP control. The outcomes of this review showed improvements in self-management behavior and medication adherence. The most successful mHealth intervention combined the feature of tailored messages, interactive communication, and multifaceted functions. Further research with longer duration and cultural adaptation is necessary. With increasing disease burden from hypertension globally, mHealth offers a potentially effective method for self-management and control of BP. mHealth can be easily integrated into existing health care systems.

Trial Registration: PROSPERO CRD42019152062; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=152062

(JMIR Mhealth Uhealth 2020;8(3):e17776) doi:10.2196/17776

KEYWORDS

hypertension; self-management; mHealth; medication adherence; mobile phone; health behavior

Introduction

Background

Hypertension is the underlying cause of around 20% of deaths globally [1]. The causes of hypertension are largely lifestyle related [2]. Poor control is known to be related to failure to diagnose cases and inadequate treatment [3]. On average, in high-income countries, around 67% of hypertension sufferers are diagnosed, with 55% treated and 28% achieving control. In contrast, in low- and middle-income countries (LMICs), around 37% of the hypertensives are diagnosed, of whom 29% are treated, with 8% achieving control [4]. The annual worldwide costs to the economy of hypertension are estimated at US $370 billion [5]. Considering unequal health care resource coverage, mobile health (mHealth) presents exciting potential in developing self-management for patients with hypertension with uncontrolled blood pressure (BP) [6].

Self-Management

The World Health Organization (WHO) defines self-management as the capability of individuals to manage their own health conditions, with or without the support of health care providers [7]. The methods of self-management include education for patients, self-monitoring of clinical data, and behavior (eg, diet, exercise, smoking, and drinking), self-titration of medical management, and support for medication adherence as per prescribed regimes [8]. A systematic review conducted by Barlow et al [9] documented that education on self-management can not only improve patients' knowledge of hypertension but can also allow for early detection of high BP. Another study reported that the self-management of hypertension enables patients to correctly self-titrate medications and irreversibly change the dynamics between doctors and patients [10]. mHealth is emerging as a vital platform to perform self-management, especially for chronic diseases [11].

Mobile Health

mHealth is the use of mobile devices—such as mobile phones, patient-monitoring devices, and wireless devices—for medical support and the delivery of health management [12]. The use of mobile devices has increased exponentially, worldwide—over 7 billion mobile device subscriptions were reported in 2015 [13]. This clearly facilitates the feasibility, generalizability, and replicability of mHealth. WHO first defined the term mHealth in 2010. Since then, mHealth technology has been more widely available and sophisticated [14,15]. mHealth employs a variety of different features, including SMS text messages, emails, phone calls, and mobile phone apps [16]. The benefits of mHealth are widely acknowledged. It can contribute to achieving universal health coverage by overcoming geographical barriers, increasing access, and the provision of health services to remote populations and underserved communities. Clinical data monitoring and educational information communications, between physicians and patients, cost less than in-person services, as the implementation of mHealth does not utilize any further resources [17]. mHealth stands at the crossroads of communication technologies and personalized health care [18].

Existing Research

Most existing systematic reviews have assessed the effectiveness of clinical outcomes of mHealth self-management in noncommunicable diseases, mainly diabetes, cardiovascular disease (CVD), and heart failure [19]. A few others have examined the content of the intervention, the study population, and economic evaluation [20,21]. Further research has investigated the effects of mHealth self-management intervention in relation to the adherence and self-titration of medication of diabetes [22]. Results were consistent in that mHealth-enabled self-management solutions could provide benefits for chronic conditions, increasing access to health care, as well as improving health care quality and patient involvement [23].

Research Gap

There is limited literature looking at the effectiveness of mHealth in hypertension self-management. A systematic review narratively synthesized the evidence for using mHealth devices to support hypertension self-management. The study reported lower systolic blood pressure (SBP) and diastolic blood pressure (DBP) in the intervention group compared with usual care [24]. Only using quantitative methods, a meta-analysis of 11 randomized controlled trials (RCTs) conducted by Lu et al [25] concluded that mHealth is an effective tool for BP control. We know of no other reviews that examined the relationship between long-term self-management and cost-effectiveness. Therefore, further exploration of the relationship between mHealth-enabled hypertension self-management and clinical outcomes needs to be conducted. This would enable relevant stakeholders to weigh the potential benefits and limitations of adopting mHealth into health care services. Thus, we aimed to determine whether mHealth is effective in improving the self-management of hypertension for adults. We systematically reviewed the existing evidence to analyze the effectiveness of mHealth-enabled self-management among hypertensive adults with the following 3 objectives:

1. 1. To measure whether the use of mHealth-enabled self-management improves the control of BP among patients with hypertension.
2. 2. To assess whether self-management education of hypertension delivered by mHealth interventions improves medication adherence and promotes lifestyle change.
3. 3. To analyze the costs of self-management support for the delivery of mHealth interventions for hypertension in adults.

Methods

Study Design

This systematic review and meta-analysis were conducted based on the original protocol (CRD42019152062) and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guideline [26].

Search Strategy

An electronic database search was conducted using PubMed, EMBASE, Web of Science, Cochrane, and Google Scholar. Searches were performed in October 2019. The key search strings consisted of 3 concepts: mHealth, hypertension, and self-management. The detailed search strategy has been presented in Textbox 1.

Textbox 1. An example of the search strategy (EMBASE).

- Search 1: (hypertension* or hypotension or hypertensive or "blood pressure"* or "elevated blood pressure" or "high blood pressure").af.
- Search 2: (self-management* or "self care"* or "self management" or "self monitoring" or self-monitoring or self-care).af.
- Search 3: (telemedicine* or telehealth or eHealth* or "e health" or e-health or mHealth* or "m health" or m-health or "mobile application" or apps or "digital health" or "mobile health" or "message text").af.
- Search 4: limit year = "2010 - 2019"
- Search 5: 1 and 2 and 3 and 4

Eligibility Criteria

We included RCTs published between January 2010 and September 2019. Only peer-reviewed articles in English were included.

Inclusion criteria were as follows: (1) adults with a primary diagnosis of hypertension; (2) the intervention used app-based tools that are accessible via mobile phone or tablet to aid self-management of hypertension; (3) reported outcomes were any one of the following: clinical data for either SBP or DBP or both, medication adherence-related outcome or change in self-management behavior.

The exclusion criteria were as follows: (1) patients with hypertension were not the main participants; (2) hypertension during pregnancy; (3) hypertension not the primary diagnosis; (4) the intervention only targeted health care providers; (5) outcomes only focused on technological development of a mobile system; (6) the mHealth service was designed for a particular target audience instead of the general public.

Observational studies, study protocols and designs, studies with abstract presentations, and duplicates were all excluded.

Study Selection

Study citations were imported and compiled into the reference management software (Endnote X8.0, Clarivate Analytics) for selection. The screening and selection of studies were conducted by 3 researchers individually (RL, NL, and FB). RL manually removed duplicates. For the initial search, RL and FB independently judged the relevance of titles and abstracts identified from electronic databases. In the second phase, FB and NL checked the study types of all remaining studies. The full text of potentially relevant articles was then retrieved. NL and RL assessed these articles against the inclusion and exclusion criteria. The snowballing method was conducted on the reference lists of relevant articles. Controversial studies and problems were compared and discussed with RL, FB, and NL.

Data Extraction

Three investigators (RL, NL, and FB) in parallel extracted the data independently and cross-checked. An adapted version of a standardized spreadsheet was used to input the data. Any disagreements were resolved through discussion. The data included the study characteristics (title, authors, year of publication, and study location); information of participants (age, gender, baseline BP, demographic information, and sample size); details about intervention and control (device, intervention and control type, message content, follow-up duration); and relevant outcome and result.

Data Synthesis and Analysis

The primary outcomes of this review were the mean SBP, DBP, and the proportion of subjects with controlled BP at the end of each trial. Patients with hypertension with SBP lower than 140 mm Hg and DBP lower than 90 mm Hg were considered to have adequate BP control [27]. Review Manager of the Cochrane Collaboration (RevMan 5.3, Cochrane Organization) was used to perform the meta-analysis.

For continuous outcomes, the effect size was defined as the mean differences (MDs) in BP between intervention and control groups. For dichotomous outcomes, the effect size was defined as the odds ratio (OR) of the proportion of patients with controlled BP between intervention and control groups. OR and MDs were derived from Manzel-Haenszel and inverse variance methods, respectively. A random-effect model was utilized to generate pooled estimates of the overall effects and reported 95% confidence intervals with all measures of effect. The I^2 statistic was used to examine inconsistencies across studies (I^2=0%-100%; more than 50% is considered as substantial statistical heterogeneity). We defined subgroups in advance to further evaluate the relationship between intervention characteristic and the clinical effect on SBP and DBP controlling for (i) frequency of reminders (tailored frequency according to the health status of participants or fixed frequency as planned); (ii) interactive patterns (interventions with a patient-provider loop interaction or without); (iii) intervention functions (single and multifaceted); and (iv) duration of trials (longer or equal to

12 months or shorter than 12 months). A sensitivity analysis was performed by excluding each study sequentially to determine the influence of any single study on the robustness of the results.

Owing to the heterogeneity in the nature of interventions and diverse outcomes of effects of self-management, we have also presented a narrative review of the findings, including structured tabular summaries according to medication adherence, change in self-management behavior, and costs.

Assessment of Risk of Bias and Quality

To account for bias and quality discrepancies, a double assessment of bias and quality was used to minimize error. RL, FB, and NL independently assessed the risk of bias and quality of the included studies. A consensus meeting was held to enable the comparison of notes from the selection of papers used within this review. An agreement was reached on conflicting points.

The risk of bias was assessed according to guidance in the *Cochrane Collaboration's Risk of Bias* handbook for RCTs [28]. Risk ratings of "low," "high," and "unclear" were assigned to each bias based on the presence of the following items: selection bias, performance bias, detection bias, attrition bias, reporting bias, and other bias. If any element was rated as high risk, the overall risk of bias was high. Funnel plots were used to detect publication bias if over 10 articles were involved in the meta-analysis. When the description of the interventions or procedures was not sufficiently detailed to judge the risk of bias, researchers contacted the study author for further information.

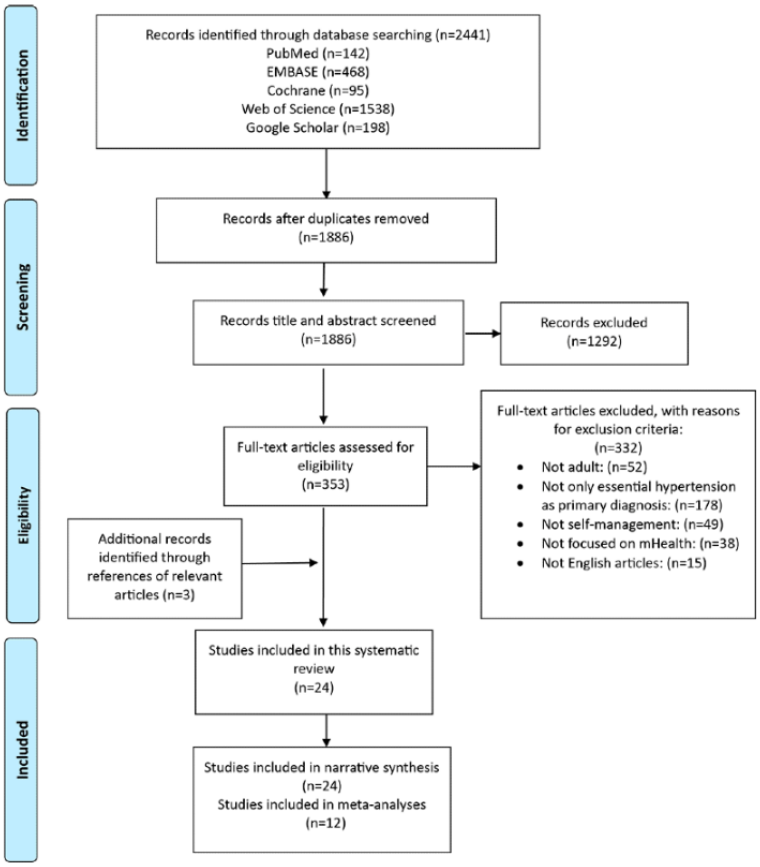
The quality of the included studies was evaluated by the Mobile Health Evidence Reporting and Assessment (mERA) Checklist [29]. This checklist was developed from the Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth (CONSORT-EHEALTH) checklist and has been extended to a wide range of mHealth interventions. It consists of 2 separate assessments: (1) essential criteria for mHealth, which comprise 16 items that identify the content, context, and technical features, to ensure the quality and generalizability of mHealth research. (2) methodological criteria, which include 26 items that are based on existing study design and study-reporting guidelines. The result was displayed by using the rate of reported items.

Results

Search Results

A total of 2441 articles were initially identified from the 5 databases; 555 articles were removed because of duplication. The remaining 1886 studies were then screened. We excluded 1292 articles because of insufficient relevance of the title and abstract to this review. A further 232 articles were rejected because they were not RCTs. A total of 9 studies were excluded as only the abstracts of protocols were available. After reviewing the full text of the remaining 353 articles, 332 studies were eliminated following the application of inclusion and exclusion criteria. An additional 3 articles were identified from references of relevant reviews, yielding a total of 24 studies. All of the 24 studies were included in narrative synthesis. Of these, 12 studies that provided SBP and DBP data and used usual care or placebo as comparisons were included in the meta-analysis. A flow diagram of the selected studies is shown in Figure 1.

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



Study Characteristics

Characteristics of the 24 studies are shown in Multimedia Appendix 1. Of which, 3 studies were from the United Kingdom, 11 from the United States, 3 from Canada, and 1 each from Spain, Taiwan (China), Chile, South Africa, Mexico, Iran, and South Korea.

Two studies were conducted in rural areas [30,31] and 22 in urban areas [32-53]. Four articles specifically targeted ethnic minorities and underserved populations [32,44]. The age range of patients across the 24 articles was 44 to 78 years. The duration of the trials ranged from 1.5 to 18 months. Half of the projects lasted for no more than 6 months. The mean sample size was 372, with a range from 54 to 1372, and 13 articles reported missing data during follow-up [32,35-46].

In all, 18 studies drew on existing theories or models to guide the intervention design [30,32,34-38,40,41,43-45,47-52]. JNC7 Guidelines and National Institute for Health and Care Excellence guidelines were most frequently used for hypertension treatment. Among 12 articles that were built on behavioral change theories [30,31,35,36,41,42,45-50], the social cognitive theory and determination theory were adopted the most. All 18 studies used treatment as usual (TAU) as a control [30-35,37-39,42-47,49,51,52]. The remaining 6 articles used the sending of different messages to the control group compared with the intervention group [36,40,41,48,50,53].

A total of 23 studies measured BP reduction [30-52]. Of which, 13 used the change in BP as their primary outcome [31-33,37-39,41-43,47,49-51] and 5 used BP control as the main outcome [33,42,44,46,52]. A total of 12 studies reported medication adherence [32-35,37,39,41,42,44,49-51]. Another 9 studies assessed the change in effectiveness of self-management behaviors [35-38,41,43,49,53]. The outcomes of self-management behaviors were varied, including readiness for behavior change, quality of life, and action plan protocol adherence. Among the 24 studies, 6 articles reported results of economic evaluation [37,42,44,46,52]. Finally, 6 articles analyzed stakeholders' satisfaction and experience with the intervention [31,32,35,37,47,50].

Intervention Characteristics

To synthesize the effects of intervention features on self-management of hypertension, this review categorized the intervention content into 13 themes: educational information of hypertension, educational information of a healthy lifestyle, self-monitoring of BP, self-monitoring of behavior change, goal setting, reminder of medication adherence, reminder of behavior change, feedback from personnel, social support, motivational encouragement, action plan, pharmacological support, and stress management. Every intervention included at least two features. Education about hypertension was included in every study. A total of 17 studies conducted BP self-monitoring [30,31,33,37,38,40-46,48,50-53]. Education about a healthy lifestyle that comprises a low-salt diet and exercise combined with goal setting in 15 studies [30,31,34-37,41,45-49,51-53]. A total of 10 studies set alerts to improve medication adherence [31,33,38,39,45-47,50-52]; 4 studies provided motivational encouragement to strengthen the patient's self-efficacy [30,43]; 3 studies provided a decision support system for an action plan

made by doctors, pharmacists, or nurses [36,48,53]; and 1 intervention included stress management [49].

A total of 6 articles included more than one intervention group [32,36,43,46,49,52,53]. Of which, 3 articles targeted the differences in effectiveness between user-driven self-management and self-management with interactive supports from personnel [32,36,46], and 1 article measured different outcomes based on varied compliance [34].

Intervention Delivery

The information about the intervention and control design was based on the intervention platform, type, and content (Multimedia Appendix 1). There were 10 interventions delivered by using SMS text messages [32,34,37,43-45,47,49-51], half of which generated automated messages and other customized messages based on the feedback of participants [32,44,45,49,50]. Six studies utilized smartphone apps [39,43,44,47,50,51] and 2 of them were interactive [47,50]. Then 6 studies reported the operation of automated emails [30,31,33,36,40,47]. Other intervention devices included wireless BP monitoring, digital medications automated or interactive voice calls, electronic medication trays, and a combination of the elements mentioned earlier.

Timing and Frequency

Intervention frequency varied considerably, and the fidelity of the intervention was not always reported. For most of the studies, a reminder to monitor BP was sent at least once a day. Most of messages about behavior or medication adherence were sent daily, and educational emails and feedbacks were sent weekly. Adherence to the intervention protocol was reported in 10 articles [30,31,35,38,39,41-45]. Six articles reported the difference between the required frequency of mHealth use and the actual use [31,33,35,45,50]. The range was from 28% to 94%.

Outcome Measures

Of the total 24 studies, 12 met the selection criteria of the meta-analysis [31-33,37-39,41-43,47,49,51]. All of them reported SBP as the primary outcome; 9 of them reported the DBP as well [33,37,38,41-43,47,49,51] and 9 articles reported the proportion of participants achieving controlled BP [31-33,35,39,41,42,44,49-51].

Meta-Analysis of Blood Pressure

A total of 3 articles had more than one intervention group with the same outcome measured [32,43,49]. Therefore, 16 interventions were shown in the forest plot of SBP, 12 interventions for DBP analysis, and 10 interventions for the comparison of BP control. As shown in Figure 2, the estimated MD of SBP between intervention and control groups was significant as −3.78 mm Hg (P<.001; 95% CI −4.67 to −2.89), with moderate heterogeneity (X^2_{15}=29.37, P=.01; I^2=49%). There was a statistically significant difference in DBP as −1.57 mm Hg (P<.001; 95% CI −2.28 to −0.86) between intervention and control groups, also shown in the forest plot with low heterogeneity (X^2_{11}=18.05, P=.08; I^2=39%; Figure 3). The OR of BP control (Figure 4) in the intervention group was 1.42 times more than that in the control group (95% CI 1.23 to 1.65). Heterogeneity was moderate (X^2_9=18.11, P=.03; I^2=50%).

Subgroup analyses were generally consistent with the main finding, showing significant reductions in SBP and DBP in the intervention groups for all subgroups analyzed (Table 1). Trials with a tailored frequency of reminders [32,37,41,42], a patient-doctor interactive loop [31,32,38,41,42,47,49], and multifaceted functions [31,37,38,42,49] showed a larger overall effect of both SBP and DBP, compared with trials with a fixed frequency of reminders [31,33,38,39,43,47,49,51], a noninteractive loop [33,37,39,43,51], and a single function [32,33,39,41,47,51]. The overall SBP reduction is greater in studies that lasted less than 12 months [31,33,37,39,49,51] than studies that lasted longer than 12 months [32,38,41,47]. Further sensitivity analyses were conducted by removing any trials sequentially, revealing no substantial difference in the overall effect for SBP and DBP. In addition, Margolis et al [42] was regarded as the main article that influenced the heterogeneity of the comparison of SBP according to the sensitivity analysis. After excluding it from the analysis, the I^2 statistic was reduced from 49% to 24% (df=15; P=.18).

Figure 2. Forest plot of the difference of systolic blood pressure between intervention and control group.

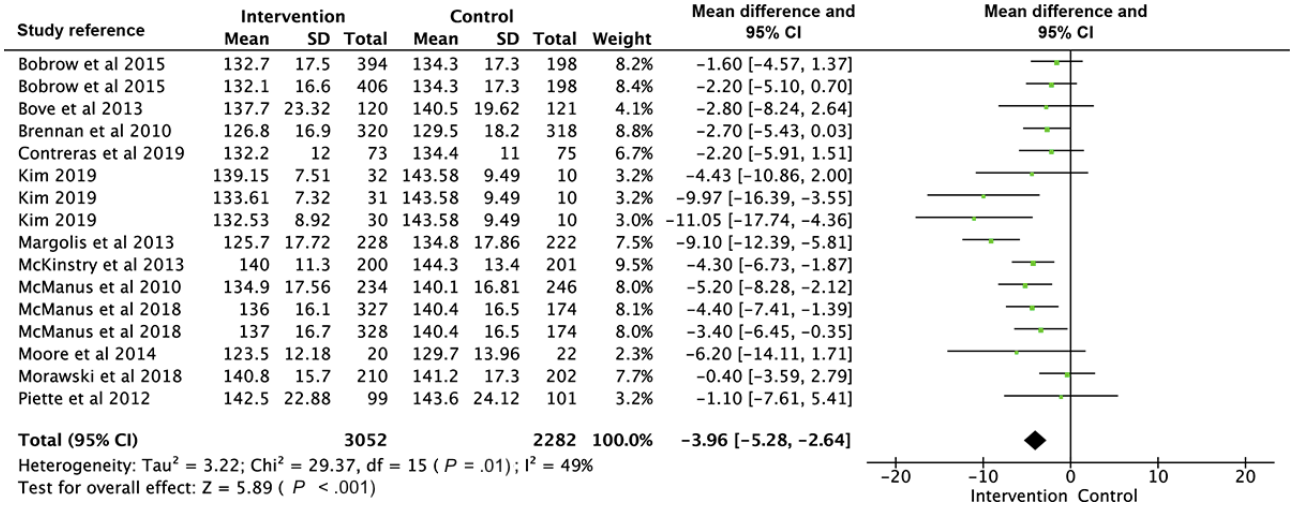


Figure 3. Forest plot of the difference of diastolic blood pressure between intervention and control group.

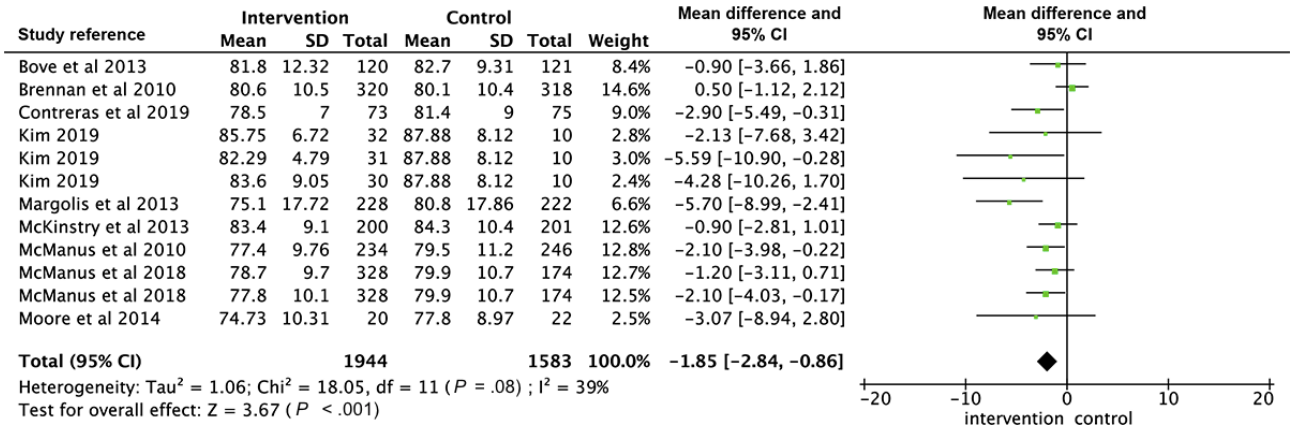


Figure 4. Forest plot of the difference of blood pressure control between intervention and control group.

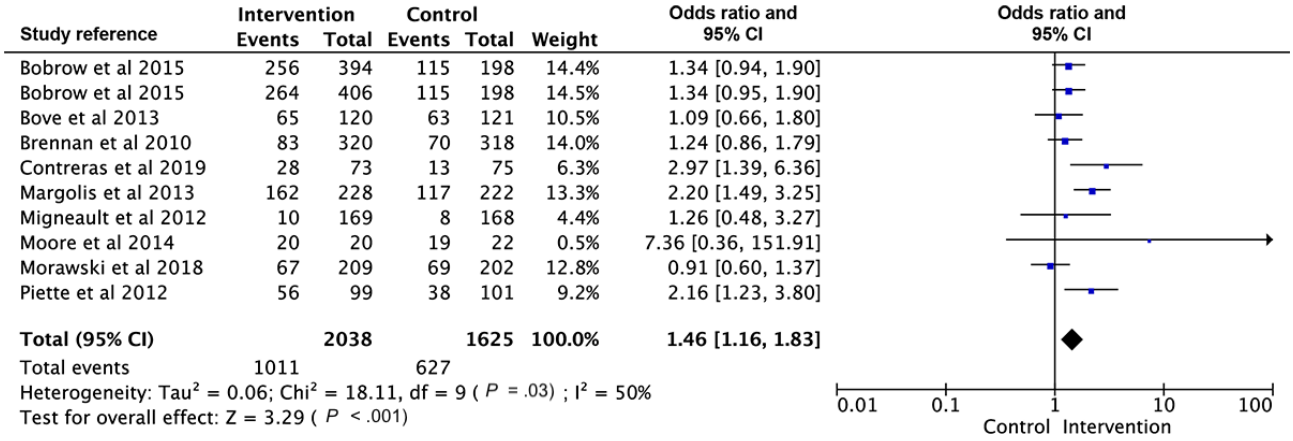


Table 1. Results from the subgroup analyses of mean differences in systolic blood pressure and diastolic blood pressure.

Analysis	Systolic blood pressure					Diastolic blood pressure				
	Studies, n	Mean difference (mm Hg)	95% CI	Heterogeneity		Studies, n	Mean difference (mm Hg)	95% CI	Heterogeneity	
				I^2 (%)	*P* value				I^2 (%)	*P* value
All studies	12	−3.78	−4.67 to −2.89	49.0	.01	9	−1.57	−2.28 to −0.86	39.0	.08
Subgroup analysis										
Reminder frequency										
Tailored frequency	4	−4.04	−6.67 to −1.41	70.0	.009	3	−2.27	−5.03 to 0.49	70.0	.04
Fixed frequency	8	−3.86	−5.41 to −2.30	36.0	.11	6	−1.70	−2.77 to −0.64	29.0	.19
Interactive pattern										
Interactive loop	7	−4.88	−7.00 to −2.75	61.0	.006	5	−2.71	−4.82 to −0.59	62.0	.01
Noninteractive loop	5	−3.17	−4.47 to −1.87	0.0	.46	4	−1.54	−2.49 to −0.59	0.0	.71
Intervention functions										
Multifaceted functions	6	−5.51	−7.25 to −3.77	44.0	.08	4	−2.20	−3.27 to −1.13	23.0	.24
Single function	6	−2.02	−3.41 to −0.63	0.0	.78	5	−0.03	−1.39 to 1.33	0.0	.40
Duration										
Shorter than 12 months	5	−4.30	−7.00 to −1.59	55.0	.04	4	−1.82	−3.06 to −0.58	0.0	.47
Longer than 12 months (12 months included)	7	−3.89	−5.44 to −2.34	50.0	.04	5	−1.85	−3.37 to −0.33	62.0	.02

Medication Adherence and Self-Management Behavior

This review narratively synthesized the outcome of medication adherence and self-management behavior (Multimedia Appendix 2). A total of 7 articles reported statistically significant improvement in medication adherence in intervention groups [32,34,39,41,49-51]. Five studies suggested that mHealth interventions improved medication adherence, despite nonsignificant outcomes [33,35,37,41,44]. Morisky Medication Adherence Scale was used in 6 studies [35,37,39,42,50,51]. As a result, Bove et al [33] reported that there is no association between medication adherence and BP control, that is, an improvement in adherence did not necessarily lead to better BP control. Of the 9 articles that focused on the behavioral change of self-management, all reported positive effects either through physical activities or through a healthier diet. Adverse events reported in studies, such as medication side-effect and cardiovascular event, were unrelated to self-management and were evenly distributed across intervention and control groups. An exception to this was McKinstry et al [37], who found 3 patients became anxious as a result of self-monitoring. Of which, 6 studies conducted qualitative research about satisfaction related to the intervention [31,32,35,37,47,50]. All showed high levels of satisfaction. Patients and physicians were keen continuing to practice mHealth.

Economic Evaluation

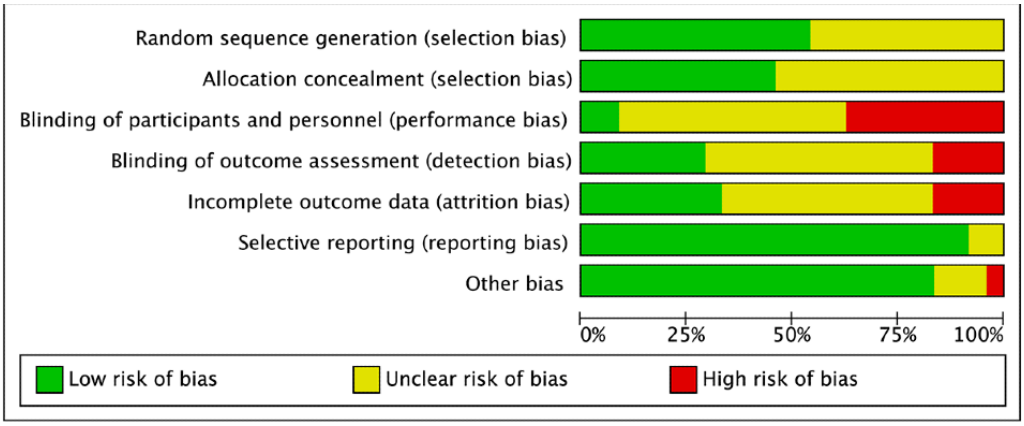
A total of 6 articles measured the cost of mHealth (Multimedia Appendix 2) [37,42,44,46,52]. In cost-saving analyses, 2 reported that the cost of mHealth interventions was higher than control [37,42]. Two studies found the cost of mHealth interventions was lower than that of control [44,47]. The measurements of expenditure varied between study settings. The main cost was from monitoring, mobile phone use, connection charges, and cost of nurse support. However, Davidson et al [44] adjusted the BP control effect into the cost analysis, which means that patients with controlled BP after receiving the experimental treatment saved the extra cost of further treatment. This showed an overall health care cost saving of over US $20,000 between intervention and control groups.

Risk of Bias

The overall risk of bias was relatively high, because no study was absolutely free of bias. Eight articles were rated as low risk for selection bias [30,35-38,48,53]. Others were judged to be unclear as the procedure of the sequence generation was not classified [31-34,38-46,50-52]. A total of 13 articles did not

describe the methods of random allocation [31,33,38-41,43-45,47,50-52]. The risk of detection bias was high in 4 articles [37,38,43,46], as these studies were unmasked to outcome assessors; 13 studies were rated as unclear, as there was an insufficient illustration of whether they blinded the outcome investigators [32-35,41,44-48,50,52]; 2 studies were double-blind trials in which participants were blinded to treatment conditions [36,53]. The control groups consisted of sending different messages compared with the intervention groups. Four articles reported no missing data from the baseline to the endpoint [47,49,51,52]. Low risk of attrition bias was found in 8 studies [43,44,46,47,49,51-53]. They reported that less than 5% of participants withdrew from the follow-up were analyses on an intention-to-treat basis for the missing data. Four studies had a high risk of attrition bias as the rate of dropout was over 15% in each study [30,33,40,48]. 92% (22/24) of the total studies were defined as low risk of reporting bias because all outcomes included in the protocol were reported in the results [30-39,41-43,45-53]. Missing pre-specified outcomes occurred in 2 articles resulting poor clarity in reporting bias [40,44]. In addition, funding bias was considered. Two articles mentioned that OMRON, which makes sphygmomanometers, including for home use, donated the BP device [31,49]. This was identified as a funding bias. Figure 5 describes the total risk of bias in the 24 studies. The funnel plot of the comparison of SBP and DBP did not show any extreme asymmetry and outliers, which suggests no significant publication bias.

Figure 5. Results of the risk of bias analysis.



Quality

According to the essential checklist of mHealth, a total of 16 items should be reported in the articles. The details of each item are demonstrated in Table 2. An average of 55% (9/16) of the 16 items was mentioned in each study, from the lowest to the highest proportion, 25% (4/16) and 81% (13/16), respectively. Among the essential items, technology platform, intervention delivery, and intervention content were reported in all 24 articles. Only 6 articles introduced the availability of infrastructure, which can support technology operations in the study site [30,31,34,41,43,47]. The rate of reporting of cost assessment was also low (6/24, 25%). The usability of the content testing, communication, and the technical solution to meet the target population were described in 8 studies [35,38,39,41,44,45,47,50], as well as the reporting rate of user feedback [30,31,35,37,38,42,47,50].

Table 2. Rate of reporting of each item in the Mobile Health Evidence Reporting and Assessment essential criteria checklist (N=24).

Items	Report rate, n (%)
Infrastructure	6 (25)
Technology platform	24 (100)
Interoperability	11 (46)
Intervention delivery	24 (100)
Intervention content	24 (100)
Content testing	8 (33)
User feedback	8 (33)
Access of individual participants	11 (46)
Cost assessment	6 (25)
Adopting input	11 (46)
Limitation for delivery at scale	13 (54)
Contextual adaptability	5 (23)
Replicability	18 (75)
Data security	12 (50)
Compliance with guideline	14 (58)
Fidelity of the intervention	10 (42)

Discussion

Summary of Principal Findings

This systematic review found 24 RCTs with 8933 adult patients with hypertension, which met the criteria to assess the effectiveness of mHealth-enabled interventions in supporting self-management. According to this meta-analysis, mHealth interventions resulted in better BP control, with a significant decrease of SBP and DBP by 3.78 mm Hg and 2.19 mm Hg, respectively, compared with usual care. All 24 studies showed a greater decrease in mHealth intervention groups than the control groups. Findings of this review confirmed that self-management education through mHealth was effective in increasing patients' knowledge of hypertension and a healthy lifestyle, medication management, and self-efficacy.

Outcomes of the economic evaluations were inconsistent across the studies. A total of 2 articles reported the negative outcomes of cost focused on direct costs [37,42]. The cost of mobile technology was shown as relatively high in rural areas. In contrast, the cost of health professionals' time in consulting in urban areas was higher than that in the rural. Thus, cost became an inevitable element when considering barriers and facilitators.

Mobile Health Intervention Design

All interventions were conducted via mobile technologies. A total of 3 elements may have contributed to the effectiveness of self-management: First, the high intensity of medication reminders. Most studies focused on medication adherence adopted weekly automated alerts and educational or motivational messages. This increased exposure to interventions, which is impossible in routine care. Brennan et al [41] conducted a comparison between the different intensities of messages and showed that a higher frequency of SMS text messages achieved better medication adherence. However, previous research has shown that reported high-dose reminders would result in response fatigue [54].

Second, user-driven designs were frequently reflected in the interventions and consisted of customized information and patient-provider loop interactions. A total of 11 studies reported 2-way communication between patients and physicians [31,32,38,41,42,46,47,49,50,52,53]. All interventions with an interactive communication loop showed significantly positive improvement in self-management behavior and BP change. These findings were consistent with the subgroup analysis in this review. Tailoring the intervention to the specific situation and readiness of patients is considered as crucial to self-management [55]. Particularly, Liu et al [36] compared the user-driven and expert-driven group behavior change. The expert-driven group showed better behavior change, perhaps because patients from the expert-driven group had more feedback, motivational commands, and support from physicians.

Finally, most of the interventions combined different functions. A total of 12 interventions had more than 2 functions [31,36-38,40,42,44,46,49,50,52,53], and 10 studies relied on SMS text messaging as their main method, while they also linked the BP monitoring devices to a Web-based system [32,34,37,43-45,47,49-51]. According to this subgroup analysis, studies with multifaceted functions had a larger effect on SBP and DBP reduction than those with a single function. In conclusion, the tailored frequency of messages based on patients' health status and readiness, two-way interactive communication, and multifaceted interventions can produce better effectiveness in the self-management of hypertension.

Strengths and Limitations of Studies Included in This Review

Significant heterogeneity showed in the meta-analysis of SBP. The reason for this is the variation in interventions. It also affects the calculation of the overall estimate [56]. According to the sensitivity analysis, Margolis et al [42] was regarded as the main article, which influenced the heterogeneity.

The occurrence of heterogeneity highlighted the strengths and limitations of the included studies. Strengths are illustrated as follows: first, more than half (13/24, 54%) of the total studies conducted power calculations for clinical data outcomes [32-38,42,43,46,48,51,53]. Second, the description of each intervention provided clear and sufficient details, allowing a thorough understanding of the method. Finally, the reporting rate of detail in the research methods was high within all articles. Over 80% (21/26) of the items listed in the mERA methodological checklist were described in these studies. This showed significant progress compared with the studies included in the previous review [57].

Particularly, self-reporting bias of compliance is clearly a potential weakness of mHealth. It would increase the risk of recall and social desirability bias. The reliability of self-reporting depends partly on the educational, socioeconomic, and cultural background of participants [58]. However, studies included in this review attempted to reduce the self-reporting bias. A total of 17 articles used self-reporting in their studies [30-35,39-46,50,52,53]. Of the 17 articles, 12 articles took steps to test the validity of self-reporting data, including home visits for behavior checks, BP monitoring devices connected to websites, and random phone calls to check medication adherence [31,33,35,39-44,47,52,53].

Referring to the limitations, the duration of the studies included was relatively short. Only 10 studies lasted for or over 1 year [32,38,41-43,46-48,51,52]. The result of subgroup analysis according to the duration of trials found in this study was similar to a previous meta-analysis, which compared digital interventions with conventional methods [59]. The overall effects of SBP and DBP were inconsistent between studies with shorter and longer durations. Thus, more evidence is needed to confirm the long-term effect of mHealth.

Though all articles were published after 2010 when the CONSORT-EHEALTH statement for reporting of eHealth and mHealth interventions was released [60], many mHealth intervention details were still unreported. Though performance bias was a prominent weakness in mHealth intervention, it can be explained by the interactive nature of the interventions, which is difficult for participants to be blinded to their health care providers [61]. Small sample size was also prominent in the included studies, which would cause a huge difference in the estimates of the target population.

In addition, all studies were from high-middle and high-income countries. Similarly, the study sites were mostly in urban settings, which restricts the diversity of the target populations. This is despite the fact that one of the important benefits of mHealth is to allow patients to receive adequate care remotely [19]. Davidson [44] reported more considerable cost savings in the mHealth group than in the control group in the study of underserved populations.

The relatively homogeneous populations limited the generalizability of the mHealth intervention. It is also important to consider culture-related differences, racial diversity, and the heterogenetic patterns of mHealth interventions, which have been mentioned in discussion of almost all articles. Nevertheless, only 5 studies have examined the potential cultural adaptation of mHealth in different settings [31,32,35,41,50]. Specifying cultural and contextual adaptabilities of mHealth interventions would help clarify whether the study design can be considered as a potentially useful platform for future research. Other observable limitations include the fact that only 6 articles reported users' satisfaction [31,32,35,37,47,50].

In relation to economic evaluations, mHealth showed only a small short-term economic benefit, but enormous potential in the longer term [62]. However, the longest duration of studies in this review is 18 months [46,52].

Strengths and Weaknesses of This Systematic Review

To our knowledge, this is the first systematic review that analyzed the relationship between the characteristics of mHealth-enabled hypertension self-management and the clinical and behavioral outcomes, using both meta-analysis and narrative synthesis. More importantly, this review adds to a body of knowledge of the strengths and limitations of included studies against the mERA checklist.

The chief weakness is the observed heterogeneities in relation to the intervention and control features. In addition, this review only recruited RCTs and excluded other designs with analyses that might also have overcome confounding. The language was restricted to English, which reduces the diversity of studies analyzed. Moreover, the sensitivity analysis was only conducted by excluding each trial sequentially to determine the influence of a single study. Owing to the small number of studies included, studies were not divided into different categories for further sensitivity analyses.

Implication of Policy Making and Further Research

Considering that at least one-third of patients with hypertension have uncontrolled BP, this review provided evidence that mHealth self-management could improve hypertension management and reduce the risks of stroke and CVD. There is increasing interest comparing benefits of mHealth approaches. Questions remain to be addressed about the values of diverse mHealth methods. To promote mHealth interventions of self-management effectively and efficiently, more clinical studies are warranted to detect the relationship between the specific intervention pattern and outcomes. In addition, patients' compliance with self-management interventions should be examined in the future.

According to the generalizability, there is a necessity to determine whether mHealth-based self-management methods should be tailored to age groups, cultural contexts, or need to be extended to include support from health care personnel. Therefore, training physicians to ensure that patients' behaviors are maintained and adopted convincingly is also necessary.

Clinical trials are called for to fill the gap of techniques of appropriate combination of mHealth intervention and routine care. Thus, more long-term economic evaluation needs to be done.

Conclusions

The intent of this systematic review was to identify and evaluate the effectiveness of mHealth-enabled self-management of hypertension from RCTs. This review clearly demonstrated that an mHealth-enabled hypertension self-management intervention was effective in improving SBP, DBP, and BP control. Both medication adherence and self-management behavior showed positive changes after the intervention. Economic evaluations presented potential cost saving in long-term effectiveness. It is the first analysis that combines clinical data and intervention features.

In conclusion, mHealth self-management has proved to be a potentially useful intervention strategy for BP management. mHealth interventions could be beneficial for BP control at the individual level and in reducing the burden of hypertension at the population level. The development of mobile technologies is especially useful when health care resources are inadequate. The broader utilization of mHealth self-management will be an important contributor to improving the quality of health care and meeting the target of universal health coverage.

Acknowledgments

This study was funded by the Center of Global Health, Zhejiang University.

Authors' Contributions

RL contributed to data search, extraction, and analysis, and then drafted and revised this paper. FB contributed to data extraction. NL advised on data analysis and revised the paper. TH initiated the research, revised the paper, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics and designs of included studies.
[DOC File , 68 KB - mhealth_v8i3e17776_app1.doc]

Multimedia Appendix 2
Outcomes of included studies.
[DOC File , 81 KB - mhealth_v8i3e17776_app2.doc]

References

1. Institute for Health Metrics and Evaluation. Global Health Data Exchange. 2016. GBD Results Tool URL: http://ghdx.healthdata.org/gbd-results-tool [accessed 2019-09-10]
2. Warburton DE, Bredin SS. Health benefits of physical activity: a systematic review of current systematic reviews. Curr Opin Cardiol 2017 Sep;32(5):541-556. [doi: 10.1097/HCO.0000000000000437] [Medline: 28708630]
3. Palafox B, McKee M, Balabanova D, AlHabib KF, Avezum AJ, Bahonar A, et al. Wealth and cardiovascular health: a cross-sectional study of wealth-related inequalities in the awareness, treatment and control of hypertension in high-, middle- and low-income countries. Int J Equity Health 2016 Dec 8;15(1):199 [FREE Full text] [doi: 10.1186/s12939-016-0478-6] [Medline: 27931255]
4. Mills KT, Bundy JD, Kelly TN, Reed JE, Kearney PM, Reynolds K, et al. Global disparities of hypertension prevalence and control: a systematic analysis of population-based studies from 90 countries. Circulation 2016 Aug 9;134(6):441-450 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.018912] [Medline: 27502908]
5. Frieden TR, Jaffe MG. Saving 100 million lives by improving global treatment of hypertension and reducing cardiovascular disease risk factors. J Clin Hypertens (Greenwich) 2018 Feb;20(2):208-211 [FREE Full text] [doi: 10.1111/jch.13195] [Medline: 29370471]
6. Morrissey EC, Glynn LG, Casey M, Walsh JC, Molloy GJ. New self-management technologies for the treatment of hypertension: general practitioners' perspectives. Fam Pract 2018 May 23;35(3):318-322. [doi: 10.1093/fampra/cmx100] [Medline: 29088438]
7. World Health Organization. 2019. Self-Care Interventions for Health URL: https://www.who.int/reproductivehealth/self-care-interventions/en/ [accessed 2019-06-10]
8. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. J Am Med Assoc 2002 Nov 20;288(19):2469-2475. [doi: 10.1001/jama.288.19.2469] [Medline: 12435261]
9. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [doi: 10.1016/s0738-3991(02)00032-0] [Medline: 12401421]

10. 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]
11. Hanlon P, Daines L, Campbell C, McKinstry B, Weller D, Pinnock H. Telehealth interventions to support self-management of long-term conditions: a systematic metareview of diabetes, heart failure, asthma, chronic obstructive pulmonary disease, and cancer. J Med Internet Res 2017 May 17;19(5):e172 [FREE Full text] [doi: 10.2196/jmir.6688] [Medline: 28526671]
12. World Health Organization. Global Diffusion of eHealth: Making Universal Health Coverage Achievable. Geneva: World Health Organization; Dec 2016.
13. International Telecommunication Union. 2015. ICT Facts & Figures URL: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2015.pdf [accessed 2019-09-02]
14. World Health Organization. Global Observatory for eHealth Series - Volume 1. Geneva: World Health Organization; Dec 22, 2010.
15. World Health Organization. mHealth - New Horizons for Health Through Mobile Technologies. Geneva: World Health Organization; 2011.
16. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
17. Wheeler TS, Vallis TM, Giacomantonio NB, Abidi SR. Feasibility and usability of an ontology-based mobile intervention for patients with hypertension. Int J Med Inform 2018 Nov;119:8-16. [doi: 10.1016/j.ijmedinf.2018.08.002] [Medline: 30342690]
18. Cameron JD, Ramaprasad A, Syn T. An ontology of and roadmap for mHealth research. Int J Med Inform 2017 Apr;100:16-25. [doi: 10.1016/j.ijmedinf.2017.01.007] [Medline: 28241934]
19. Chaet D, Clearfield R, Sabin JE, Skimming K, Council on Ethical and Judicial Affairs American Medical Association. Ethical practice in telehealth and telemedicine. J Gen Intern Med 2017 Oct;32(10):1136-1140 [FREE Full text] [doi: 10.1007/s11606-017-4082-2] [Medline: 28653233]
20. Chib A, van Velthoven MH, Car J. mHealth adoption in low-resource environments: a review of the use of mobile healthcare in developing countries. J Health Commun 2015;20(1):4-34. [doi: 10.1080/10810730.2013.864735] [Medline: 24673171]
21. Wang V, Smith VA, Bosworth HB, Oddone EZ, Olsen MK, McCant F, et al. Economic evaluation of telephone self-management interventions for blood pressure control. Am Heart J 2012 Jun;163(6):980-986. [doi: 10.1016/j.ahj.2012.03.016] [Medline: 22709750]
22. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-Management education and support. J Diabetes Sci Technol 2017 Sep;11(5):1015-1027 [FREE Full text] [doi: 10.1177/1932296817713506] [Medline: 28560898]
23. So CF, Chung JW. Telehealth for diabetes self-management in primary healthcare: A systematic review and meta-analysis. J Telemed Telecare 2018 Jun;24(5):356-364. [doi: 10.1177/1357633X17700552] [Medline: 28463033]
24. Alessa T, Abdi S, Hawley MS, de Witte L. Mobile apps to support the self-management of hypertension: systematic review of effectiveness, Usability, and User Satisfaction. JMIR Mhealth Uhealth 2018 Jul 23;6(7):e10723 [FREE Full text] [doi: 10.2196/10723] [Medline: 30037787]
25. Lu X, Yang H, Xia X, Lu X, Lin J, Liu F, et al. Interactive mobile health intervention and blood pressure management in adults. Hypertension 2019 Sep;74(3):697-704. [doi: 10.1161/HYPERTENSIONAHA.119.13273] [Medline: 31327259]
26. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. Br Med J 2009 Jul 21;339:b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
27. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Himmelfarb CD, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol 2018 May 15;71(19):e127-e248 [FREE Full text] [doi: 10.1016/j.jacc.2017.11.006] [Medline: 29146535]
28. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
29. 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. Br Med J 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
30. Nolan RP, Liu S, Shoemaker JK, Hachinski V, Lynn H, Mikulis DJ, et al. Therapeutic benefit of internet-based lifestyle counselling for hypertension. Can J Cardiol 2012 May;28(3):390-396. [doi: 10.1016/j.cjca.2012.02.012] [Medline: 22498181]
31. Piette JD, Datwani H, Gaudioso S, Foster SM, Westphal J, Perry W, et al. Hypertension management using mobile technology and home blood pressure monitoring: results of a randomized trial in two low/middle-income countries. Telemed J E Health 2012 Oct;18(8):613-620 [FREE Full text] [doi: 10.1089/tmj.2011.0271] [Medline: 23061642]

32. 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 (SMS-text adherence support [StAR]): a single-blind, randomized trial. Circulation 2016 Feb 9;133(6):592-600 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.017530] [Medline: 26769742]
33. Bove AA, Homko CJ, Santamore WP, Kashem M, Kerper M, Elliott DJ. Managing hypertension in urban underserved subjects using telemedicine--a clinical trial. Am Heart J 2013 Apr;165(4):615-621. [doi: 10.1016/j.ahj.2013.01.004] [Medline: 23537980]
34. Varleta P, Acevedo M, Akel C, Salinas C, Navarrete C, García A, et al. Mobile phone text messaging improves antihypertensive drug adherence in the community. J Clin Hypertens (Greenwich) 2017 Dec;19(12):1276-1284 [FREE Full text] [doi: 10.1111/jch.13098] [Medline: 28941056]
35. Migneault JP, Dedier JJ, Wright JA, Heeren T, Campbell MK, Morisky DE, et al. A culturally adapted telecommunication system to improve physical activity, diet quality, and medication adherence among hypertensive African-Americans: a randomized controlled trial. Ann Behav Med 2012 Feb;43(1):62-73. [doi: 10.1007/s12160-011-9319-4] [Medline: 22246660]
36. Liu S, Brooks D, Thomas SG, Eysenbach G, Nolan RP. Effectiveness of user- and expert-driven web-based hypertension programs: an RCT. Am J Prev Med 2018 Apr;54(4):576-583. [doi: 10.1016/j.amepre.2018.01.009] [Medline: 29456025]
37. McKinstry B, Hanley J, Wild S, Pagliari C, Paterson M, Lewis S, et al. Telemonitoring based service redesign for the management of uncontrolled hypertension: multicentre randomised controlled trial. Br Med J 2013 May 24;346:f3030 [FREE Full text] [doi: 10.1136/bmj.f3030] [Medline: 23709583]
38. McManus RJ, Mant J, Bray EP, Holder R, Jones MI, Greenfield S, et al. Telemonitoring and self-management in the control of hypertension (TASMINH2): a randomised controlled trial. Lancet 2010 Jul 17;376(9736):163-172. [doi: 10.1016/S0140-6736(10)60964-6] [Medline: 20619448]
39. Morawski K, Ghazinouri R, Krumme A, Lauffenburger JC, Lu Z, Durfee E, et al. Association of a smartphone application with medication adherence and blood pressure control: the MedISAFE-BP randomized clinical trial. JAMA Intern Med 2018 Jun 1;178(6):802-809 [FREE Full text] [doi: 10.1001/jamainternmed.2018.0447] [Medline: 29710289]
40. Lee P, Liu J, Hsieh M, Hao W, Tseng Y, Liu S, et al. Cloud-based BP system integrated with CPOE improves self-management of the hypertensive patients: A randomized controlled trial. Comput Methods Programs Biomed 2016 Aug;132:105-113. [doi: 10.1016/j.cmpb.2016.04.003] [Medline: 27282232]
41. Brennan T, Spettell C, Villagra V, Ofili E, McMahill-Walraven C, Lowy EJ, et al. Disease management to promote blood pressure control among African Americans. Popul Health Manag 2010 Apr;13(2):65-72 [FREE Full text] [doi: 10.1089/pop.2009.0019] [Medline: 20415618]
42. Margolis KL, Asche SE, Bergdall AR, Dehmer SP, Groen SE, Kadrmas HM, et al. Effect of home blood pressure telemonitoring and pharmacist management on blood pressure control: a cluster randomized clinical trial. J Am Med Assoc 2013 Jul 3;310(1):46-56 [FREE Full text] [doi: 10.1001/jama.2013.6549] [Medline: 23821088]
43. McManus RJ, Mant J, Franssen M, Nickless A, Schwartz C, Hodgkinson J, TASMINH4 investigators. Efficacy of self-monitored blood pressure, with or without telemonitoring, for titration of antihypertensive medication (TASMINH4): an unmasked randomised controlled trial. Lancet 2018 Mar 10;391(10124):949-959 [FREE Full text] [doi: 10.1016/S0140-6736(18)30309-X] [Medline: 29499873]
44. Davidson TM, McGillicuddy J, Mueller M, Brunner-Jackson B, Favella A, Anderson A, et al. Evaluation of an mHealth medication regimen self-management program for African American and Hispanic uncontrolled hypertensives. J Pers Med 2015 Nov 17;5(4):389-405 [FREE Full text] [doi: 10.3390/jpm5040389] [Medline: 26593951]
45. Meurer WJ, Dome M, Brown D, Delemos D, Oska S, Gorom V, et al. Feasibility of emergency department-initiated, mobile health blood pressure intervention: an exploratory, randomized clinical trial. Acad Emerg Med 2019 May;26(5):517-527. [doi: 10.1111/acem.13691] [Medline: 30659702]
46. Bosworth HB, Powers BJ, Olsen MK, McCant F, Grubber J, Smith V, et al. Home blood pressure management and improved blood pressure control: results from a randomized controlled trial. Arch Intern Med 2011 Jul 11;171(13):1173-1180. [doi: 10.1001/archinternmed.2011.276] [Medline: 21747013]
47. Moore JO, Marshall MA, Judge DC, Moss FH, Gilroy SJ, Crocker JB, et al. Technology-supported apprenticeship in the management of chronic disease: a randomized controlled trial. J Clin Outcomes Manag 2014 Mar;21(3):110-122 [FREE Full text]
48. Nolan RP, Feldman R, Dawes M, Kaczorowski J, Lynn H, Barr SI, et al. Randomized controlled trial of e-counseling for hypertension: REACH. Circ Cardiovasc Qual Outcomes 2018 Jul;11(7):e004420. [doi: 10.1161/CIRCOUTCOMES.117.004420] [Medline: 30006474]
49. Kim M. Effects of customized long-message service and phone-based health-coaching on elderly people with hypertension. Iran J Public Health 2019 Apr;48(4):655-663 [FREE Full text] [Medline: 31110975]
50. Chandler J, Sox L, Kellam K, Feder L, Nemeth L, Treiber F. Impact of a culturally tailored mHealth medication regimen self-management program upon blood pressure among hypertensive Hispanic adults. Int J Environ Res Public Health 2019 Apr 6;16(7):pii: E1226 [FREE Full text] [doi: 10.3390/ijerph16071226] [Medline: 30959858]
51. Contreras EM, Rivero SM, García ER, López-García-Ramos L, Vilas JC, Suárez AB, Compliance Group of Spanish Society of Hypertension (SEH-LELHA). Specific hypertension smartphone application to improve medication adherence in

hypertension: a cluster-randomized trial. Curr Med Res Opin 2019 Jan;35(1):167-173. [doi: 10.1080/03007995.2018.1549026] [Medline: 30431384]
52. Maciejewski ML, Bosworth HB, Olsen MK, Smith VA, Edelman D, Powers BJ, et al. Do the benefits of participation in a hypertension self-management trial persist after patients resume usual care? Circ Cardiovasc Qual Outcomes 2014 Mar;7(2):269-275. [doi: 10.1161/CIRCOUTCOMES.113.000309] [Medline: 24619321]
53. Ghezeljeh TN, Sharifian S, Isfahani MN, Haghani H. Comparing the effects of education using telephone follow-up and smartphone-based social networking follow-up on self-management behaviors among patients with hypertension. Contemp Nurse 2018;54(4-5):362-373. [doi: 10.1080/10376178.2018.1441730] [Medline: 29451091]
54. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. PLoS One 2014;9(2):e88166 [FREE Full text] [doi: 10.1371/journal.pone.0088166] [Medline: 24505411]
55. Lavielle M, Puyraimond-Zemmour D, Romand X, Gossec L, Senbel E, Pouplin S, et al. Methods to improve medication adherence in patients with chronic inflammatory rheumatic diseases: a systematic literature review. RMD Open 2018;4(2):e000684 [FREE Full text] [doi: 10.1136/rmdopen-2018-000684] [Medline: 30116556]
56. Sedgwick P. Meta-analyses: what is heterogeneity? Br Med J 2015 Mar 16;350:h1435. [doi: 10.1136/bmj.h1435] [Medline: 25778910]
57. Mileski M, Kruse CS, Catalani J, Haderer T. Adopting telemedicine for the self-management of hypertension: systematic review. JMIR Med Inform 2017 Oct 24;5(4):e41 [FREE Full text] [doi: 10.2196/medinform.6603] [Medline: 29066424]
58. Garber MC, Nau DP, Erickson SR, Aikens JE, Lawrence JB. The concordance of self-report with other measures of medication adherence: a summary of the literature. Med Care 2004 Jul;42(7):649-652. [doi: 10.1097/01.mlr.0000129496.05898.02] [Medline: 15213489]
59. McLean G, Band R, Saunderson K, Hanlon P, Murray E, Little P, DIPSS co-investigators. Digital interventions to promote self-management in adults with hypertension systematic review and meta-analysis. J Hypertens 2016 Apr;34(4):600-612 [FREE Full text] [doi: 10.1097/HJH.0000000000000859] [Medline: 26845284]
60. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
61. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst Rev 2012 Dec 12;12:CD007459 [FREE Full text] [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]
62. Stoddart A, Hanley J, Wild S, Pagliari C, Paterson M, Lewis S, et al. Telemonitoring-based service redesign for the management of uncontrolled hypertension (HITS): cost and cost-effectiveness analysis of a randomised controlled trial. BMJ Open 2013 May 28;3(5):pii: e002681 [FREE Full text] [doi: 10.1136/bmjopen-2013-002681] [Medline: 23793650]

Abbreviations

BP: blood pressure
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth
CVD: cardiovascular disease
DBP: diastolic blood pressure
LMIC: low- and middle-income country
MD: mean difference
mERA: Mobile Health Evidence Reporting and Assessment
mHealth: mobile health
OR: odds ratio
RCT: randomized controlled trial
SBP: systolic blood pressure
TAU: treatment as usual
WHO: World Health Organization

Edited by G Eysenbach; submitted 12.01.20; peer-reviewed by T Davidson, L de Witte; comments to author 31.01.20; revised version received 11.02.20; accepted 26.02.20; published 27.03.20.

Please cite as:
Li R, Liang N, Bu F, Hesketh T
The Effectiveness of Self-Management of Hypertension in Adults Using Mobile Health: Systematic Review and Meta-Analysis
JMIR Mhealth Uhealth 2020;8(3):e17776
URL: http://mhealth.jmir.org/2020/3/e17776/
doi:10.2196/17776
PMID:32217503

Original Paper

Prevention of HIV and Other Sexually Transmitted Infections by Geofencing and Contextualized Messages With a Gamified App, UBESAFE: Design and Creation Study

Felipe Besoain[1,2*], PhD; Antoni Perez-Navarro[2,3*], PhD; Constanza Jacques Aviñó[4*], MPH, PhD; Joan A Caylà[5*], MD; Nicolas A Barriga[1], PhD; Patricia Garcia de Olalla[4,6], MD, MPH

[1]School of Videogames Development and Virtual Reality Engineering, Faculty of Engineering, Universidad de Talca, Campus Talca, Chile

[2]Internet Interdisciplinary Institute, Universitat Oberta de Catalunya, Castelldefels, Spain

[3]Faculty of Computer Sciences, Multimedia and Telecommunication, Universitat Oberta de Catalunya, Barcelona, Spain

[4]Epidemiologic Service - Public Health Agency of Barcelona, Plaça Lesseps 1, Barcelona, Spain

[5]Foundation of the Tuberculosis Research Unit of Barcelona, Barcelona, Spain

[6]Biomedical Research Network on Epidemiology and Public Health (CIBEResp), Barcelona, Spain

[*]these authors contributed equally

Corresponding Author:
Antoni Perez-Navarro, PhD
Faculty of Computer Sciences, Multimedia and Telecommunication
Universitat Oberta de Catalunya
Rambla del Poblenou, 156
Barcelona, 08018
Spain
Phone: 34 619963691
Email: aperezn@uoc.edu

Abstract

Background: Advances in the development of information and communication technologies have facilitated social and sexual interrelationships, thanks to the websites and apps created to this end. However, these resources can also encourage sexual contacts without appropriate preventive measures in relation to HIV and other sexually transmitted infections (STIs). How can users be helped to benefit from the advantages of these apps while keeping in mind those preventive measures?

Objective: This study aimed to prevent STIs by helping users to remember preventive measures in the risky situations.

Methods: We have used the design and creation methodology and have developed a software system. This system has two parts: an Android operating system app with emphasis on ubiquitous computing and gamification as well as a server with a webpage. First, a functional test with 5 men who have sex with men (MSM) allowed us to test the app with end users. In addition, a feasibility test with 4 MSM for a month allowed us to try the UBESAFE system with all its functionalities.

Results: The main output is a system called UBESAFE that is addressed to MSM. The system has two main parts: (1) an app that sends preventive contextualized messages to users when they use a contact app or when they are near a point where sexual contacts are likely and (2) a server part that was managed by the public health agency of Barcelona (ASPB), which preserves the quality and pertinence of messages and places and offers instant help to users. To increase users' adherence, UBESAFE uses a gamified system to engage users in the creation of preventive messages. Users increased the initial pool of messages by more than 100% (34/30) and created more than 56% (9/16) of places (named hot zones).

Conclusions: The system helped MSM who used it to become conscious about HIV and other STIs. The system also helped the ASPB to stay in contact with MSM and to detect behaviors that could benefit from preventive measures. All functions were performed in a nonintrusive manner because users used the app privately. Furthermore, the system has shown how important it is to make users a part of the creation process as well as to develop apps that work by themselves and thus become useful to the users.

(JMIR Mhealth Uhealth 2020;8(3):e14568) doi:10.2196/14568

KEYWORDS

human immunodeficiency virus; mobile apps; sexually transmitted infection; recreational games

Introduction

Background

Currently, many health organizations work actively to decrease the number of HIV infections. Despite the major advances in the treatment of HIV, prevention of infection is still better than treatment [1]. In the recent years, the number of health campaigns and the number of locations where such health campaigns are implemented have grown substantially [2]. Nowadays, many different types of methods are available to disseminate information about health and prevention of diseases. Some of the main methods include publicity projects, outreach work with groups of individuals who may be at risk, the monitoring and control at a national level of items recognized as having a negative impact on health, programs at educational institutions, and the use of social media, to name a few [3]. Nevertheless, it is important to consider the impact of the information and communication technology in social relationships. Many studies show how the internet has been considered as a connection point to meet sexual partners [4,5].

However, with the introduction of smartphones and tablets to the market, information is increasingly omnipresent and, as a consequence, the access to apps and social networks is more ubiquitous [6,7]. Dating mobile apps are also modern tools to find sexual partners through the internet, considering aspects such as location, timing, and taste, among others. It has also been shown how dating apps impact HIV infection for many reasons: first, users can find sexual partners easily on the go; second, users maximize the likelihood to find a sexual partner because apps are free; and third, there are dating apps for heterosexual and gay and bisexual communities. People who use the dating apps and have these sexual conducts usually take more risks when they have a sexual encounter [8].

Researchers have examined the main tools that are currently used for preventing the spread of HIV [9]. These tools include Web-based and mobile apps, games, and social media, among others. The target of many of these campaigns and of several strategic plans to control HIV infections are men who have sex with men (MSM) because they are considered a high-risk group in most European and American countries. In this context, most efforts have been directed toward prevention through the use of apps that provide information when facing specific situations and behavior. It has also been suggested that researchers in public health should work with app developers to incorporate innovative elements, starting with interventions that reduce the risk and the associated behaviors, as well as that improve the inclusivity and interactivity of the apps [10].

Choi et al [11] showed how an app can help introduce healthier behaviors regarding HIV, although they found no concluding remarks about risk reduction. Alarcon et al [12] have shown that sending messages through apps helps to promote testing for HIV and other sexually transmitted infections (STIs). Biello et al [13] proposed a study to analyze an app that promotes the uptake of HIV testing and pre-exposure prophylaxis. Although this last app has a similar target to that proposed in this paper, it is an informative app that is focused on informing about how to get tested and where to get prophylaxis measures, but not on sending context-based messages as the one proposed in this paper. Chow et al [14] show how using geolocalized apps offer several opportunities in HIV prevention. However, as far as we know, this is the first time that context of the users is used to prevent risky behaviors in HIV, thanks to the characteristics and omnipresence of smartphones.

This Work

This work addresses these drawbacks by developing an app to send preventive notifications to users when it detects situations such as the activation of a particular app (dating app) on their smartphone, or their proximity to areas with a high probability of intercourse. To increase adherence, the app uses gamification techniques. The development process has been performed in a co-design process with potential users (MSM) with a goal of developing a system that has value for users per se [15] and increases users' adherence and preventive effect. This work is the continuation of work by Besoain et al [16] in which the use of mobile devices and their ubiquity was used to prevent STIs.

We based our design of this technological approach on the elaboration likelihood model that describes a framework of multiple processes in which communication variables (eg, channel or message) can change people's attitudes and ultimately their behavior [17]. Several investigations in social psychology have consistently shown that the thoughts that people generate in response to social information are important predictors of their attitudes and behaviors [18]. Many of the studies on attitude change use specific communicative information (persuasive messages) to generate thoughts of different directionality. Therefore, messages have been used with arguments in a favorable or unfavorable direction (see for a review [19]). When a person receives arguments that are strong, this tends to generate thoughts in line with the information [20].

Therefore, in this app, participants received, and were asked to help the health community to generate, favorable arguments for healthy behavior, such as using condoms, with the aim of generating thoughts in this direction and positively impacting users' future decisions. However, it is important to note that in this study, our objective was to develop the technology and test the main idea in general terms. More specific studies about the effects of the arguments for healthy behavior with this technology will be a part of future work.

This paper describes an app developed for HIV prevention called UBESAFE. To prevent users from feeling that the app has an overbearing monitoring effect, it has been designed in a way that allows them to actively participate in the creation of a tool to reinforce healthy behavior. Users download UBESAFE and configure it themselves. They can use the app in two main ways, to reinforce healthy behavior when (1) using contact apps that could be used for initiating sexual relationships (Grindr, ManHunt, etc) and (2) walking or passing by a geographical

area (hotzone) where sexual activities could occur, for example, gay saunas and nightclubs. The users select the contact apps alone or the contact apps along with geographical areas that they want to include in UBESAFE, taking an active role in the process. When users engage in one of these activities (use a preselected contact app or move through a preselected hotzone), UBESAFE sends a health notification message. These HIV prevention messages are not designed to discourage sexual relationships, but rather to encourage users to make healthy decisions and increase the awareness of their sexual health (for example, do not forget to use condoms). The HIV prevention messages are written by MSM with the guidance of health professionals. Furthermore, to encourage an active role, users can write their own messages. In a gamification aspect of the app, users can earn points if they post messages. It is important to note that UBESAFE does not intend to stigmatize sex, but rather encourage awareness and healthy sex decisions to contribute to HIV prevention.

The paper is structured as follows: (1) the methods used to create the app are introduced, and the architecture and use cases of the app are shown; (2) the results, which are mainly the main features of the app and its utility as a preventive tool, are presented; and (3) our conclusions and future work in this area are described.

Methods

Overview

The method used in this research follows the *design and creation* approach to create a system with two parts: a management web server for health professionals; and an app for users, which is the main preventive element. The app uses ubiquitous computing concepts [21], localization services [22], and game mechanics [23,24] to prevent HIV infections and other STIs. This process has been divided in two steps: (1) a co-design methodology to enhance and address the functional requirements of the app by a group of potential users [15], and (2) a functional and a feasibility test has been performed to evaluate the app with a target group [25,26].

To design and develop the app, the following methodologies have been followed:

- The design process was performed with health experts and the target group. It is important to note that because the app will be a supervised app, both health experts and target group will be users of different parts of the app, and they both will pursue different goals. This co-design methodology plays a key role because we plan to create an app useful for both kinds of users, and it is important to create something with added value for every single user, but trying to avoid interference, from the user experience (UX) point of view, between the different goals. The co-design was performed through focus groups, surveys, and functional testing. The result is an incremental process that has gone through three previous apps:
  - Ubiapp used geofencing to recognize different hotspots (places with high probability of sexual encounter) and the use of a list of risky apps (dating apps) for delivering a health message [16]. In the development of this app, a study with 17 MSM helped obtain a pool of messages that were later used as the seed pool of messages of UBESAFE.
  - Ubinut launches health messages addressed to prevent obesity and allow the users to score them on a Likert scale [27].
  - Geonut has the same functionality as Ubinut, but it also incorporates geofencing to hotspots (in this case, places with restaurants, fast food courts, among others).

To test the app and choose the health or preventive messages, a test was performed in two steps. First, a functional test was performed with 5 MSM for 2 weeks with the aim of receiving feedback on the UX through a focus group. The number of the participants in the study was based on the open call for volunteers through the public health agency of Barcelona (ASPB); the volunteers were MSM HIV-negative. The number is considered sufficient because the objective of this study was to test technical aspects of the app in a real situation and also receive initial feedback from the user. The second study was a feasibility test that was then performed with 4 MSM for a month to try the UBESAFE system with all of its functionalities. These participants were the same as those from the first study, except for one participant who could not continue because of lack of time. The purpose of this initial test was to receive feedback on the UX through a focus group with MSM. The previous experiences with Ubiapp, Ubinut, and Geonut were taken into account to test the system.

The UBESAFE system has two major components (Figure 1):

1. Simple web interface system (SWIS), which allows the health administrator to add, modify, and delete messages and point of interest (POI, known as hot zone)
2. Mobile app developed for Android operating system version 6.0.1 allows the mobile users to receive notifications based on a smart context in three ways:
   - Browsing a hotlink (URL Patrol)
   - Using a risk app
   - Being nearby a hot zone (Localization)

These notifications are any of both health messages provided by a health administrator and user private messages. User can also interact with other functionalities such as the following:

- Evaluate a message with a Likert scale
- Share a message or POI with the web server
- See user stats on the gamification module
- Contact a health professional

Figure 1. Architecture of solution and components of the UBESAFE system. This includes (1) simple web interface system, (2) web server, and (3) Android clients. SWIS: simple web interface system.

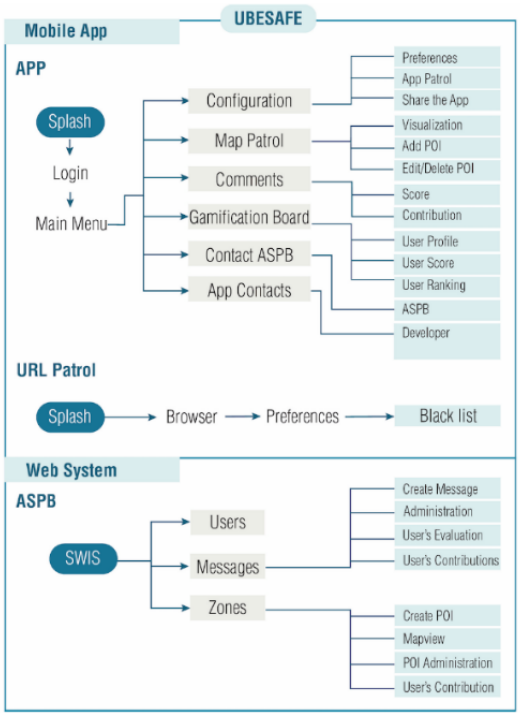
Flux of Work of UBESAFE System

To understand all functionalities and how the app is structured, the following system is presented (see Figure 2):

1. Mobile app: the mobile app has two main activities:
   - URL patrol, which has all the functionalities of a web browser but with the preventive system incorporated.
   - UBESAFE, which works by detecting the different situations that can make the users aware of their actions through a health message. Thus, the message will be ranked by the users on a Likert scale from 1 to 5, where 1 is the least interesting and 5 is the most interesting.
2. SWIS, through which the health administrator is able to approve, modify, or delete the messages and POIs shared by the users. The system also shows statistics of users' scoring and most valued messages, number of users, etc.

As can be seen from Figure 1, it is important to note that UBESAFE is an app supervised by health professionals, who can benefit from the knowledge taken from the app to improve their prevention campaigns and also ensure that the messages in the app will always be true and respectful.

Figure 2. Flux of work of the app. ASPB: public health agency of Barcelona; SWIS: simple web interface system; POI: point of interest.



UBESAFE App

The mobile app UBESAFE has two main activities with their own functionalities:

- First, UBESAFE includes the detection of any of both proximity to specific areas (what we call *hot zones*) and the use of any apps that the user wants to be warned on using it (what we call *risky apps*).
- Second, URL patrol notifies the users when it detects what we call a *hot URL*, that is, a URL that the user has marked as one to be warned when clicking on it.

It is important to note that all health messages are retrieved from a local database. This database is controlled from the SWIS and updated every time that a new message or POI is detected, and the administrator releases a new version of the database.

The functionalities and the flux of work for UBESAFE is presented in Figure 2.

The first time that UBESAFE is run on the mobile device, the mobile users will have to fill in their data and configure the app. The process is performed in three steps. First, once the app is opened by touching the UBESAFE icon, the app will check the user data. Second, a preference list will be shown to fill with their information. Finally, once the users have entered all the information, the app will process in background to sign in the users to the SWIS database and download the messages and POI available from the SWIS to query them locally (this process happens on the login section of the Figure 2). Once the mobile users have completed the information requested by the app, the health administrator will be able to see the users' data in the users' section of the SWIS system, as will be shown in SWIS section. The app allows the users to participate in the community sharing data for research purposes, or they can choose to be anonymous.

Detecting Risk Apps

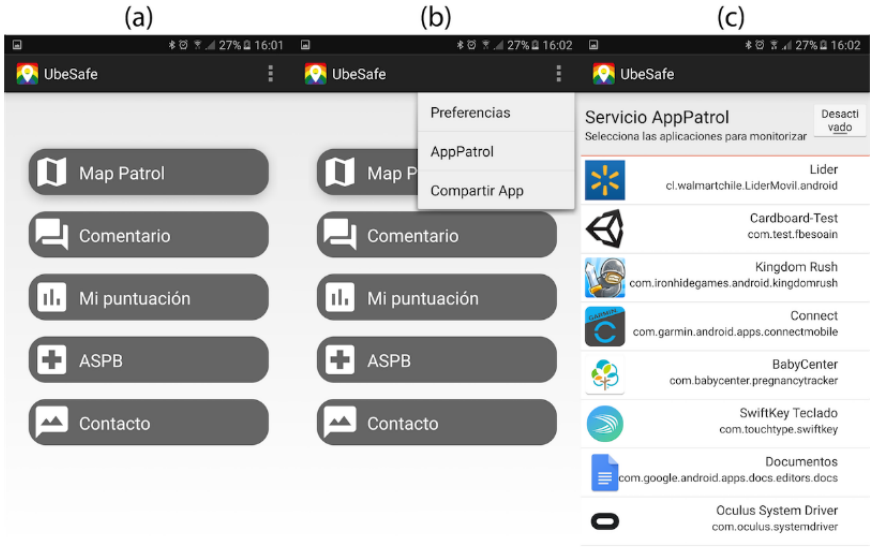
For detecting risk apps, the users need to configure the service known as AppPatrol. This service will show the mobile users all the apps installed on their device. The mobile users will select the apps for monitoring and then activate the service. This service, similar to all the services of UBESAFE, works in background, and the mobile users do not need to start it again. It will continuously be monitoring the device until the service is deactivated.

It is important to note that it is the user who decides to add the option to receive a warning on using the app. If the users do not add, they will not get any message on using the app. It is important to note that all the process is within the smartphone, and no data are stored regarding the use of the app or the apps for which the user has added to be warned.

Figure 3 shows three states of the app at different times: (1) once the app is opened by touching the UBESAFE icon, (2) the users select in the preference section the AppPatrol settings, and (3) a list with all the icons and name of the installed apps will be shown (this service is off by default).

Once the users activate the service, they can choose the apps they wish to be warned about. Right part of Figure 3 shows the list of installed apps, the user selects apps from the list to monitor. This action is performed by doing a long press on the list (according to the mobile standards, a long action present selection on a list). Thereafter, the user starts the service. This service is always on (algorithms to optimize this service are shown in section Detecting Hot Zones).

Figure 3. UBESAFE—Android client: setting the AppPatrol service on the app.



Detecting Hot Zones

The module to detect hot zones is known as Map Patrol (see Figure 4); it is responsible for sending health messages to users when the app detects the proximity to areas with a high probability of intercourse (hot zones). In addition, the UX has been enhanced with several functionalities:

- To add private POI to the database of the mobile devices.
- To share the POI with the community sending the information to the Information server.

- To delete any POI from the database, allowing the mobile users to choose which hot zones to detect.

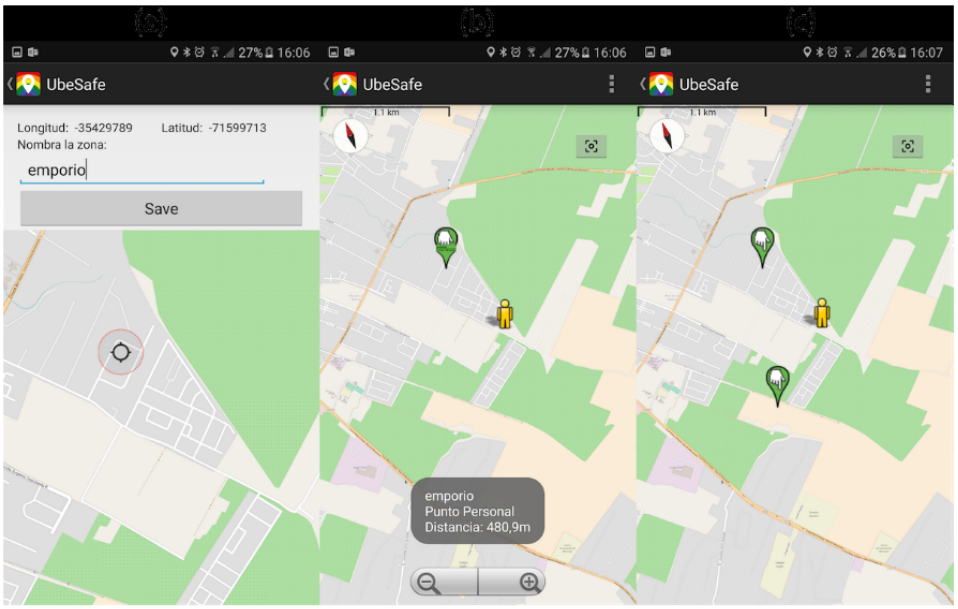
When the user has selected the Map Patrol option, the app opens a mapView with the user's current positions and POI or hot zones nearby. In the configuration section, the mobile users can add POIs and manage a single POI (share it with the community or delete it). It is important to note that UBESAFE allows users to add their own POI, and users themselves decide if they want to share them with the community or not. This option helps to increase the value of the app for users because they can use it to keep their own POI.

Figure 4 shows three states of the app at different times:

- First, the users select adding POI in the Map Patrol settings. In the mapView interface, the users select from a mapView a point by doing a long press on the map. The app automatically will get the latitude and longitude; thereafter, the users must write a name of the POI.
- Second, the app shows the mapView with the recently added POI. If the user touches the POI, then UBESAFE shows information and distance from the user's current position to the POI.
- Third, on the right, two POIs can be seen nearby the user's current position.

Besides adding their own POI to the app, mobile users also have the option of sharing the POI with the community by sending the information to the information server, and deleting any POI from the database, allowing the mobile users to choose which hot zones to detect. To do this, (1) the users have to select to manage POI in the Map patrol settings; (2) the app shows a list with the POIs (it is important to note that each POI has an icon to the left that shows the current status of the POI: shared or local); and (3) by doing a long press on the target POI, the mobile users can delete the POI from the local database. In addition, the users can select the POI to share with the community by doing a simple press.

Figure 4. UBESAFE—Android client: adding users' favorite hot zones for future monitoring of the alert service.



Managing Health Messages

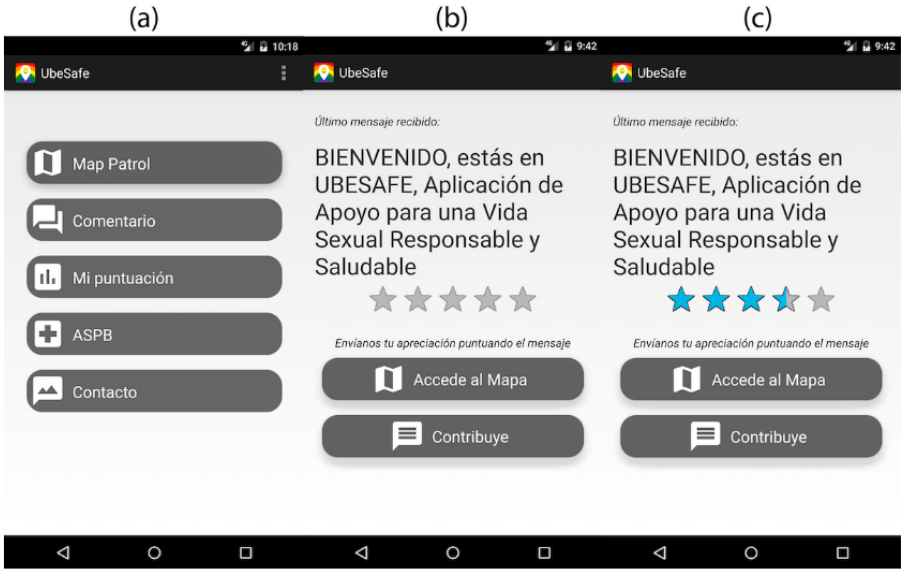
To improve the functionality and UX related to the health messages, UBESAFE allows the mobile users to score the health message received. The app uses the notification manager service of the Android operating system and presents the message on the notification bar. Moreover, because the app is running in the background all the time, there is no action needed by the users to activate this detection.

This interface includes two shortcuts to the map patrol; thus, the mobile users can see their position and all the POIs nearby, which allows the mobile users to add, delete, and upload their own health messages. The mobile users can access this interface from the main menu in *Comments* (see left side of Figure 5) or any time they get a new notification.

When the users open the notification, the app will prompt them to rank the received message out of 5 stars, as it is shown in the center and right side of Figure 5. After the ranking process, the app sends this score to the SWIS where the health administrator can see the average and the highest scored message.

Mobile users can add their own messages because they can have private messages for their consideration. They can also, if they wish, upload, and share their private messages with the community. When the users want to share a message, it will be uploaded to the information server where the health administrator can review it through the SWIS. This revision could modify the original message if needed. Thereafter, the message will be added to the system main database. Thus, users can contribute and add the messages that they think can be more helpful to result in a behavioral change.

Figure 5. UBESAFE—Android client: primary interface for scoring messages with shortcuts for contribution and mapView.



To share a message with the community, the users must perform the following procedure: (1) the app shows a list with the messages, and the mobile users can add or upload a message to the information server; (2) the users, by doing a simple press, select the message and then select the sync symbol; and (3) the message has changed the icon from the left, showing that the private message is updated to the web server.

Gamification Scoreboard

UBESAFE has a scoreboard with the most valued health messages and users' ranking. Every time that mobile users share a POI or health message, they get an amount of points of experience in exchange: the amount of experience awarded is related to the number of actions that the users have performed in the system; the more messages or POIs they share with the community, the more experience they get. Depending on the amount of experience, mobile users will get a medal that reflects their rank in the system. The contribution will also be presented on a scoreboard, enhancing the experience with the system and promoting the sharing and contributing of POI and messages to the UBESAFE system.

In the gamification section of the app, first, the app shows the main menu, where the users select the punctuation; and second, the users can see the number of shared contributions (messages and POI). They can also see a bar of experience and the current medal; the image from the right shows how the users have increased their experience in the bar. This happens because the mobile users have shared more messages and POIs.

The use of experience points, medals, and ranking in UBESAFE is an example of gamification because it uses these elements to motivate users to participate more actively in the community. It promotes a sense of competition between the users and allows them to see their accomplishments.

It is important to note that gamification is within the part of prevention and has two extra effects: (1) users get involved in the preventive messages, and messages can be better tuned for target users preferences; and (2) by sharing POIs and messages, health service can improve the design of prevention campaigns by the language used, as well as by the places in which to launch campaigns.

URL Patrol UBESAFE

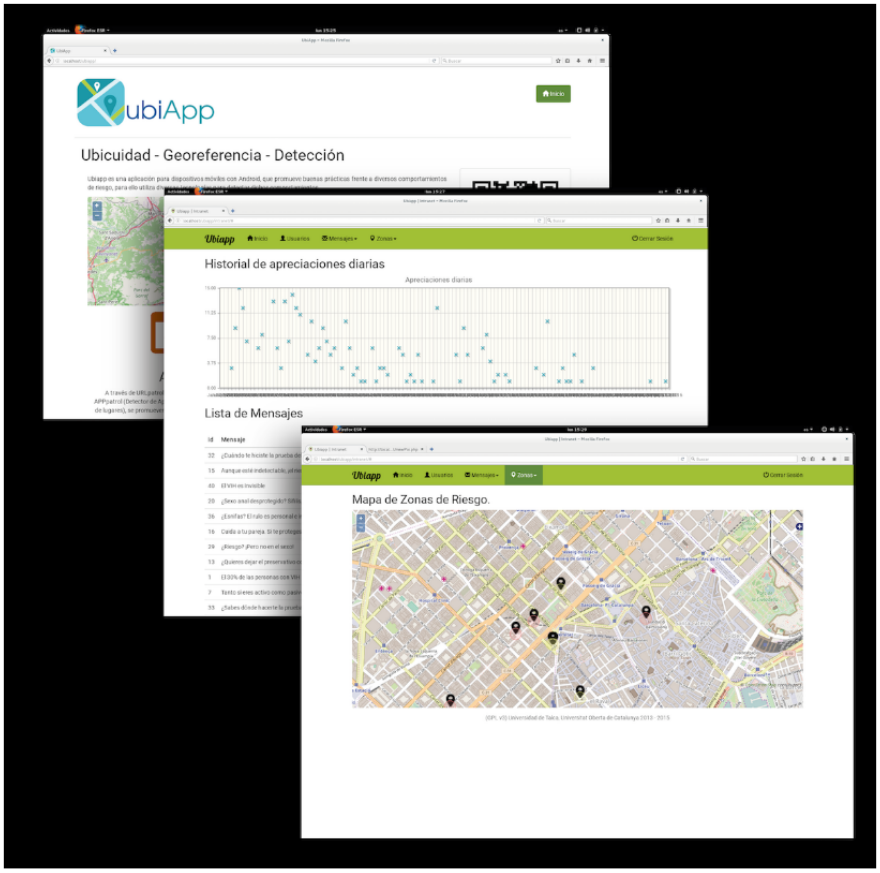
URL Patrol is part of the UBESAFE app, but it could be run independently as a web browser because it allows users to navigate on the internet. They can open the URL patrol preferences, where they can add or delete any website that users want to be warned about when getting into it. The kind of URLs that one can expect to find here are those related to contact apps or contact websites. The app comes with a preloaded list of websites such as Grindr [28], Manhunt [29], Tinder [24,30].

When the users navigate on a website that is on the list, the system will detect that action and will notify the mobile user with a health message. It is important to mention that URL Patrol is not another mobile app, but part of UBESAFE and can be run independently for UX purposes. Mobile users can configure URL patrol as their default web client and use it to navigate on the internet as part of the detection and prevention system.

Simple Web Information System

The SWIS is the interface that allows a health professional to review and check the messages and POIs shared by the mobile users. This interface is important because it is part of a workflow that is controlled by a health professional. The workflow secures the information and validates the messages that will be sent to the mobile users. Figure 6 represents the SWIS with their modules and functionalities.

Figure 6. UBESAFE—web client: web interface where the health professional can access and manipulate all the data related to the UBESAFE system.



Once the health administrator gets into the SWIS, the following modules will be presented:

- Users: information about the users (user name, nationality, age, and UID, the internal code that identifies the mobile devices where the UBESAFE app was installed).
- Messages: here the health administrator will be able to create, manage, and delete messages from the system. This will also include the messages shared by the users.
  - Create message: the administrator can add a health message to the system. The message can be in three languages (English, Spanish, and Catalan). If the administrator does not include the message in one of the available languages, the message will not appear in that language.
  - Administration: the administrator can update or delete a specific message. This option is used to modify mobile users' contributions or refine a proposed health message.
  - User's evaluation: in this section, the administrator can see the frequency of scored messages by the mobile users per day (graph). Moreover, a list is shown ordered by the average scoring of the system for each message, and the frequency of the scoring by the users per day.
  - User's contribution: a list with the messages shared by the users is shown. The administrator can update the message, translate it, and approve or delete it. Once the message is approved, it is considered for the next update of the database and shared to the users' community.
- Zones: here, the health administrator will be able to create, manage, and delete hot zones from the system. This will also include the hot zones and POIs shared by the users.
  - Create POI: the administrator, through positioning a POI into the map interface, is able to add a hot zone to the database.
  - mapView: here the hot zones are shown. Thus, the administrator can have a global and geographic perspective of the data that are stored into the system.
  - POI administration: the administrator can see, approve, or delete POIs in the system. The interface provides the latitude and longitude information and also can show the point on a map, by using Google Maps interface.
  - User's contribution: a list with the POIs contributed by the users is shown. The administrator can see, approve, or delete every single POI.

See Figure 2 for flux of work of the SWIS.

Optimization and Key Points

Optimization is a key factor for mobile devices because the main power source (the battery) is limited. The resources that spend more battery on a mobile device are the screen, GPS, long processing times, and an internet connection. Taking this into consideration, all the algorithms were optimized to provide the maximum efficiency in the use of resources.

Optimization and refining have been done during the whole process of software development, considering all the apps

developed. In UBESAFE, *AlarmManager* and *broadcaster receiver* allow the app to control the different states of the mobile device during its uptime.

For the alert service, the following states were considered to save battery.

1. It is detected when the mobile device is connected to a USB cable. The connection could be done for two purposes:
   - To charge the mobile device
   - To connect to a computer

In both cases, the mobile device is generally not being used outside, and therefore, there is no need to use the detection of hot zones.

1. The algorithm was enhanced for localization, polling the GPS less than the alert service of the other mobile apps with high accuracy.
2. The update service is not running all the time. It is responsible for retrieving data from the information server every other day because the versions of data are expected to change in a period of days.

For the risk service, the following premise was considered to save battery. Every time that the users are using a contact app on their smartphones, they are using the device. Therefore, the app with the risk services will only be detected when the screen is on. Otherwise, it is assumed that the smartphone is off or in a standby mode. Hence, the current state of the device is checked, using the *PowerManager* of Android application programming interface. Thus, it is possible to infer what the users are doing with the device, starting the risk service when it is necessary rather than all the time.

Finally, the AlarmManager provides access to system-level alarm services. Using the AlarmManager allows an app to schedule tasks that may need to run or repeat beyond the scope of its lifecycle. The Android system tries to batch alarms at similar intervals or times together to preserve battery life. By batching alarms from multiple apps, the system can avoid frequent device waking and networking. Therefore, it saves battery and resources.

Ethics and Consent Statement

We did not have to ask for ethical approval because at the time of the study, it was not legally necessary to ask for ethical approval for a study where no health-relevant information and personal data were collected from participants because they were already registered in the ASPB, and no personal data are stored within the app. Nevertheless, the users volunteered to become part of the project, received no financial compensation, and could leave the project whenever they wanted. In addition, all the information was shared with them.

Test the App

As mentioned in the overview, two tests were performed on the app. The first was a functional test with the objective to receive user feedback about technical aspects and usability, whereas the second was a feasibility test to try the UBESAFE system with all of its functionalities. Below, the results are discussed in further detail.

Results

Functional Testing Results

The demographic characteristics of the sample are shown in Table 1. Two of the participants are in their 30s and two are older than 40 years. Four of them were Spanish, and one was from Chile.

All the volunteers declared using the mobile device as their primary device for accessing the internet. Moreover, all five of them have an internet plan on their devices. Therefore, they were fully connected the whole time. In this context, they also declared to be constantly aware if the devices have some notification or message. This effect was increased by the ubiquity of the information. Today, a notification on the mobile device is information from a message, text, email, or game notification, among others.

Table 1. Sociodemographic characteristics of volunteers.

Age (years)	Occupation	Level of studies	Country of birth
27	Nurse	Undergraduate degree	Spain
42	Interior designer	Undergraduate degree	Spain
30	Medic	Undergraduate degree	Chile
37	Receptionist	Professional degree	Spain
45	Actor	Undergraduate degree	Spain

In addition, all five of the volunteers mentioned that they used the mobile device to chat with friends. Meanwhile, two of them also declared that they use the mobile device to chat with unknown people. In this context, the mobile device was also used to search for information related to health topics such as STIs, HIV, health centers, and sports.

The volunteers used UBESAFE for 2 weeks with the following detections:

- *Execution of some kind of app*: applies in situations where users open apps designed for contacting sexual partners, such as *Manhunt*.
- *Proximity to a geographical zone where sexual contacts often take place*: applies in situations where users enter or are near to what we call a *hot zone*.
- *Detection of a target URL*: applies in situations where users open a target URL.

It is important to note that this version of the app only sends a health message. Once the notification has been displayed, the software can interact with other installed apps, for example, allowing the users to share the notification through email, text message, or social networks. Thus, users have a fully connected experience that can also help to promote prevention among others. When they receive a notification, the users can make an informed decision regarding the possible consequences of their behavior. As a result, the use of this software raises users' awareness of their actions and encourages them to take steps to limit the spread of STIs.

All of the volunteers declared that they were able to install and configure the app without problems and the app did not compromise the standard functioning of the devices. Regarding the three main functionalities of UBESAFE in this version (detection of hot zones, risk apps, and URL of contact), two of the volunteers mentioned that they received health notification when they were near a hot zone or using a risk app, and three received a health message through the URL detection when they were navigating on the internet.

As part of the discussion and conclusions of the experience of the focus group, Table 2 describes the highlights to be considered: with this first initial testing, it was possible to test the three most important functionalities of detection of the apps and know the users' perceptions with the aim of enhancing the UX of the system. This test also had a technical tracking of bugs and issues through the Google Play platform for developers. Finally, there is a continuous refinement of the modules of the app as part of the iterative development methodology. All these experiences have increased the value of the product, with more emphasis on the users than on the process.

Table 2. Perceptions of volunteers in the functional test.

Volunteers comments	Analysis
I got a lot of messages for the use of one application	This feature was developed on purpose in the first version to see how the mobile users will react to the notifications on demand. Next versions allowed a configuration of the timing of the notifications.
Battery consumption when I use the map	Battery consumption is an issue in all apps that require the constant processing of data. In this case, the use of the mapView consumes energy from two principal sources: localization and screen.
Repetition of the health messages in the different detections	This happens because the experiment had 20 messages for testing purposes. The messages will be presented randomly. A big database of health message is required to avoid repetition.
It is necessary to have more hot zones	Some zones are provided by the health administrator, but there is knowledge that only MSM[a] know and could be beneficial for the community and future interventions of the public health service.

[a]MSM: men who have sex with men.

Feasibility Testing Context

Four MSM volunteers were enrolled to participate actively in the feasibility evaluation of the UBESAFE system. They were enrolled in Barcelona, a city with an important offer of gay leisure. All the participants declared that they utilized social networks through their smartphones frequently (at least twice a day). They were aged between 27 and 45 years. It is important to consider that the UBESAFE system aims to have a preventive role, making users more conscious of their actions.

Data Sources and Collection

The data were collected through the smartphones owned by the users. All the data obtained from the system were stored in the database as the foundation of the analysis. Moreover, the data were collected through the SWIS and analyzed concurrently. It is relevant to mention that all the users agreed to share the information for further analysis; setting this option was a requirement to run the app.

Data were collected at three different times: first, through the entry survey; second, through the smartphones; finally, through a final focus group. The objective of the entry survey was to learn more about the target group, regarding their knowledge, habits, and behaviors. The aim of the final focus group was to learn about the UX experience with UBESAFE and collect software suggestions.

The app collects the following three types of data at different times:

1. Profile information: this information is collected the first time that the users run the app. At that moment, the users' contact information is requested such as user name, email, age, and nationality. In addition, the users must set if they want to share this information with the system or use the system anonymously. With these data, users' profiles are created on the SWIS that later are related with their scores on the health messages.
2. Perceptions: when UBESAFE detects any of the three detection systems, it will notify the users in the notification bar of the smartphone. The users will be prompted to read the health message and score it on a scale from 1 to 5 (Figure 5).
3. Question for the health administrator: the users were able to send questions to the health administrator in charge (professional of the ASPB).

The health administrator through the SWIS can access to see and review all the data from the system. Several actions can be taken related to the Create, Read, Update, and Delete (CRUD) of messages and POIs on the system. Moreover, the health administrator has a key role in the revision and approval of health data provided by the smartphone users. In addition to the management options, the SWIS shows the list of messages and graphs of perceptions of the users. These features add value to the system because the administrator of the system can see in real time the positive or negative impact of a health message.

Feasibility Evaluation

In the feasibility evaluation, 1 health professional from the ASPB participated, writing the initial database with 30 messages and responding to the users' questions. The initial messages were taken from the study presented in the UBIAPP testing experiment, and the questions were responded to within 24 hours. The study had a duration of 30 days; 357 evaluation of health messages were registered in total in the system (2.9 evaluation average per volunteer per day).

The messages come from two sources:

1. Health administrator: the test was begun with 30 messages related to prevention of risky behaviors in MSM. Ten messages were taken from the UBIAPP study. In that study, volunteers were asked to choose the messages they found most and least suitable as preventive messages and were also offered the chance to propose new messages. In addition to this initial list of messages, the ASPB provided 20 additional messages that were reviewed by health professionals of the institution. The messages were displayed randomly by the app.
2. Smartphone users: users are able to share messages and POI with the system. Therefore, a sustainable way was created to maintain the system with new information and encourage users through a gamification system. To keep the primary objectives of the messages, POI, and system, the health administrator must review these data.

The mean of the ratings received during the whole period was 4.60, indicating that the information sent to participants was highly rated in general. A total of 64 messages were registered in the SWIS; therefore, 34 messages were shared by users and added to the 30 original pool of messages. Ten more messages were shared and were not considered appropriate to be added to the pool.

It is important to highlight those messages scored with the highest rating by the participants, considering only those with scores above average plus one standard deviation. Four out of 10 of the highest rated messages were shared by the users rather than the administrator. These messages are shown in Table 3.

Now that the information about the health messages has been seen, it is necessary to turn to the POIs and hot zones. The initial database started with seven hot zones entered by the health administrator. After the test, the database ended up with 16 hot zones. Therefore, 56% (9/16) of the hot zones were contributions of the smartphone users. This information is very important for the health service because it allowed them to identify new zones for launching prevention campaigns.

Table 3. Messages sent by the system scored with the highest rating by the participants.

Message	Average scoring	Frequency
When was the last time that you got tested for HIV?	5.0	10
Even if nothing has been detected, 0 risk doesn't exist!	4.95	10
HIV is invisible	4.91	11
Unprotected anal sex? You can get syphilis, gonorrhea and other STI[a]	4.86	11
Do you snort? The tube is personal and non-transferable	4.80	10
Take care of your partner. If you protect yourself, you protect him	4.79	12
Risk? But not in sex!	4.70	10
Do you want to stop using a condom with your guy? Let's get tested together	4.70	15
Oral sex has also some risks	4.68	10
Let us be serious against HIV.	4.6	9

[a]STI: sexually transmitted infection.

Discussion

In this paper, we presented the app UBESAFE, which sends preventive notifications to users when it detects situations such as the activation of particular apps on their smartphones, the access to a specific URL on the internet, or their proximity to areas with a high probability of intercourse (*hot zones*). It also develops community for the users, considering their ideas and knowledge of the hot zones. To provide a sustainable way of getting new data, the main experience was developed with *gamification* concepts. It is important to note that UBESAFE wants to create awareness of each detected situation through health messages, considering privacy and users' preferences.

Therefore, the app works on four main lines:

- Sending health messages to users when the app detects a hot URL, that is, a website where users can meet or chat with unknown people.
- Sending health messages to users when the app detects the use of a contact app (that we call risk apps such as Manhunt, Tinder, Badoo, or Brenda, among others).
- Sending health messages to users when the app detects the proximity to areas with a high probability of intercourse (hot zone), such as saunas, intercruising zones, etc.
- Allowing the users to make a community sharing messages and POI (hot zones) through the system, enhancing the experience with a gamified scoreboard.

UBESAFE adds new modules and functionalities to the previously developed Android apps Ubiapp [12], Ubinut, and Geonut [19] through its modular architecture. It is a supervised

app designed not only to have a health system that helps to keep information within the app respectful and true but also to help these professionals to know their target better.

The app was tested for 30 days. During the trial period, 357 evaluations were received from users, which rated the different messages on a Likert scale of 1 to 5, obtaining an average of nearly 2.9 responses a day per user. The mean of the ratings received during the whole period was 4.60, indicating that the information sent to participants was highly rated in general.

The volunteers highly valued the functionalities related to sharing information and seeing how their peers valued it. They felt they are part of an informed community that helps to improve the knowledge on this matter, enhancing the UX and purpose behind installing and being part of this app. In fact, users increased the pool of messages by more than 100% (34/30) and created 56% (9/16) of the hot zones.

The app shows, on one hand, its value to the target users, and how the gamification can increase adherence to the app, and on the other hand, how it can help the health professionals to know their target users and prevention campaigns better. Thus, UBESAFE is an app that can help in the prevention of HIV and other STIs.

Finally, further work will address the following:

1. Improving the graphic design and UX feedback of the gamified board
2. Implementing a more accurate ontology algorithm to improve the automatic system that sends messages in UBESAFE, making the selected messages closer to those that a health professional would choose in each situation
3. Measuring the impact of UBESAFE as a ubiquitous system for promoting healthy habits through an evaluation methodology, such a longitudinal study.
4. Explore a theoretical perspective from the social psychology of communication to advance an explanation of the effects of STI prevention strategies through mobile devices, using a theory-based technological solution that implies the development of positive attitudes toward the use of condoms and positive sexual health behaviors through an active generation of consequent favorable thoughts (ie, the Elaboration Likelihood Model of persuasion [17]).

Acknowledgments

The authors would like to thank the Epidemiology Service of the ASBP.

Conflicts of Interest

None declared.

References

1. World Health Organization.: World Health Organization; 2016. Global Health Sector Strategy on HIV 2016-2021. Towards Ending AIDS URL: https://www.who.int/hiv/strategy2016-2021/ghss-hiv/en/ [accessed 2020-01-20]
2. UNAIDS. 2019. Global HIV & AIDS Statistics — 2019 Fact Sheet URL: https://www.unaids.org/en/resources/fact-sheet [accessed 2020-01-20]
3. Simas C, Larson H, Schulz W, Kamel-Boulos M, Apfel F, Gauci AJ, et al. Utilising Social Media to Support HIV/STI Prevention: Evidence to Inform a Handbook for Public Health Programme Managers. Stockholm, Sweden: European Centre for Disease Prevention and Control (ECDC); Jul 2017.
4. Cooper A, Morahan-Martin J, Mathy RM, Maheu M. Toward an increased understanding of user demographics in online sexual activities. J Sex Marital Ther 2002;28(2):105-129. [doi: 10.1080/009262302528518861] [Medline: 11894795]
5. Daneback K, Månsson SA, Ross M. Using the internet to find offline sex partners. Cyberpsychol Behav 2007 Feb;10(1):100-107. [doi: 10.1089/cpb.2006.9986] [Medline: 17305455]
6. Shim R. Tablets impact the notebook market: enter the ultrabook. Inf Disp 2012;28(2-3):12-14. [doi: 10.1002/j.2637-496x.2012.tb00471.x]
7. Martin C. TheINQUIRER.: The Inquirer; 2012 Jan 4. 73 Million Tablets Were Shifted in 2011 URL: https://www.theinquirer.net/inquirer/news/2135275/million-tablets-shifted-2011 [accessed 2020-01-20]
8. Choi EP, Wong JY, Lo HH, Wong W, Chio JH, Fong DY. The impacts of using smartphone dating applications on sexual risk behaviours in college students in Hong Kong. PLoS One 2016;11(11):e0165394 [FREE Full text] [doi: 10.1371/journal.pone.0165394] [Medline: 27828997]
9. Levine D. Using technology, new media, and mobile for sexual and reproductive health. Sex Res Soc Policy 2011;8(1):18-26. [doi: 10.1007/s13178-011-0040-7]
10. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J Med Internet Res 2013 Jan 4;15(1):e1 [FREE Full text] [doi: 10.2196/jmir.2301] [Medline: 23291245]
11. Chiou P, Liao P, Liu C, Hsu Y. Effects of mobile health on HIV risk reduction for men who have sex with men. AIDS Care 2019 Sep 26:1-9. [doi: 10.1080/09540121.2019.1668531] [Medline: 31558040]

12. Gutiérrez MA, Quevedo MF, Valle SM, Jacques-Aviñó C, David ED, Caylà JA, et al. Acceptability and effectiveness of using mobile applications to promote HIV and other STI testing among men who have sex with men in Barcelona, Spain. Sex Transm Infect 2018 Sep;94(6):443-448. [doi: 10.1136/sextrans-2017-053348] [Medline: 29626174]
13. Biello KB, Marrow E, Mimiaga MJ, Sullivan P, Hightow-Weidman L, Mayer KH. A mobile-based app (MyChoices) to increase uptake of HIV testing and pre-exposure prophylaxis by young men who have sex with men: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2019 Jan 7;8(1):e10694 [FREE Full text] [doi: 10.2196/10694] [Medline: 30617042]
14. Chow JY, Klausner JD. Use of geosocial networking applications to reach men who have sex with men: progress and opportunities for improvement. Sex Transm Infect 2018 Sep;94(6):396-397. [doi: 10.1136/sextrans-2018-053622] [Medline: 29930157]
15. Mirri S, Roccetti M, Salomoni P. Collaborative design of software applications: the role of users. Hum Cent Comput Inf Sci 2018 Mar 19;8(1). [doi: 10.1186/s13673-018-0129-6]
16. Besoain F, Perez-Navarro A, Caylà JA, Aviñó CJ, de Olalla PG. Prevention of sexually transmitted infections using mobile devices and ubiquitous computing. Int J Health Geogr 2015 May 3;14:18 [FREE Full text] [doi: 10.1186/s12942-015-0010-z] [Medline: 25935151]
17. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. Adv Exp Soc Psychol 1986;19:123-205 [FREE Full text] [doi: 10.1016/S0065-2601(08)60214-2]
18. Eagly AH, Chaiken S. The advantages of an inclusive definition of attitude. Soc Cogn 2007;25(5):582-602. [doi: 10.1521/soco.2007.25.5.582]
19. Petty RE, Wegener DT. Attitude change: multiple roles for persuasion variables. In: Lindzey G, Fiske ST, Gilbert DT, editors. The Handbook of Social Psychology. Oxford, UK: Oxford University Press; 1998.
20. de Miguel J, Gallardo I, Horcajo J, Becerra A, Aguilar P, Briñol P. The impact of stress on processing persuasive messages. Rev Psicol Soc 2014;24(3):399-409. [doi: 10.1174/021347409789050597]
21. Oates BJ. Researching Information Systems and Computing. London: Sage Publications Ltd; 2006.
22. Android Developers. Localize Your App URL: https://developer.android.com/guide/topics/resources/localization [accessed 2020-01-20]
23. Deterding S, Dixon D, Khaled R, Nacke L. From Game Design Elements to Gamefulness: Defining 'Gamification'. In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek'11; September 28 - 30, 2011; Tampere, Finland URL: https://dl.acm.org/doi/10.1145/2181037.2181040 [doi: 10.1145/2181037.2181040]
24. Basten D. Gamification. IEEE Softw 2017;34(5):76-81. [doi: 10.1109/ms.2017.3571581]
25. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
26. 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]
27. Besoain F, Perez-Navarro A, Ojeda F, Reyes-Suarez JA. Promoting Healthy Nutrition Behavior Using Mobile Devices and Ubiquitous Computing. In: Proceedings of the International Work-Conference on Ambient Assisted Living. 2015 Presented at: IWAAL'15; December 1-4, 2015; Puerto Varas, Chile p. 89-100 URL: https://link.springer.com/chapter/10.1007%2F978-3-319-26410-3_9 [doi: 10.1007/978-3-319-26410-3_9]
28. Grindr - Find Gay, Bi, Curious Guys for Free Near You. URL: http://grindr.com/ [accessed 2020-01-20]
29. Manhunt. Gay Dating, Chat and Hookups. URL: http://www.manhunt.net/
30. Tinder | Match. Chat. Date. URL: https://tinder.com/ [accessed 2020-01-20]

Abbreviations

ASPB: public health agency of Barcelona
MSM: men who have sex with men
POI: point of interest
STI: sexually transmitted infection
SWIS: simple web interface system
UX: user experience

Edited by G Eysenbach; submitted 07.05.19; peer-reviewed by H Sanchez, C Granell, C Grov; comments to author 29.09.19; revised version received 22.11.19; accepted 29.11.19; published 17.03.20.

Please cite as:
Besoain F, Perez-Navarro A, Jacques Aviñó C, Caylà JA, Barriga NA, Garcia de Olalla P
Prevention of HIV and Other Sexually Transmitted Infections by Geofencing and Contextualized Messages With a Gamified App, UBESAFE: Design and Creation Study
JMIR Mhealth Uhealth 2020;8(3):e14568
URL: http://mhealth.jmir.org/2020/3/e14568/
doi:10.2196/14568
PMID:32181752

Original Paper

Mobile Assessment of Acute Effects of Marijuana on Cognitive Functioning in Young Adults: Observational Study

Tammy Chung[1], PhD; Sang Won Bae[2], PhD; Eun-Young Mun[3], PhD; Brian Suffoletto[4], MD; Yuuki Nishiyama[5], PhD; Serim Jang[6]; Anind K Dey[7], PhD

[1]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

[2]School of Systems and Enterprises, Stevens Institute of Technology, Hoboken, NJ, United States

[3]Department of Health Behavior and Health Systems, University of North Texas Health Science Center, Fort Worth, TX, United States

[4]Department of Emergency Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

[5]Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

[6]Dietrich College of Humanities and Social Sciences, Carnegie Mellon University, Pittsburgh, PA, United States

[7]Information School, University of Washington, Seattle, WA, United States

Corresponding Author:
Tammy Chung, PhD
Department of Psychiatry
University of Pittsburgh School of Medicine
3811 O'Hara Street
Pittsburgh, PA, 15213
United States
Phone: 1 4123524651
Email: tammychung111@gmail.com

Abstract

Background: Mobile assessment of the effects of acute marijuana on cognitive functioning in the natural environment would provide an ecologically valid measure of the impacts of marijuana use on daily functioning.

Objective: This study aimed to examine the association of reported acute subjective marijuana high (rated 0-10) with performance on 3 mobile cognitive tasks measuring visuospatial working memory (Flowers task), attentional bias to marijuana-related cues (marijuana Stroop), and information processing and psychomotor speed (digit symbol substitution task [DSST]). The effect of distraction as a moderator of the association between the rating of subjective marijuana high and task performance (ie, reaction time and number of correct responses) was explored.

Methods: Young adults (aged 18-25 years; 37/60, 62% female) who reported marijuana use at least twice per week were recruited through advertisements and a participant registry in Pittsburgh, Pennsylvania. Phone surveys and mobile cognitive tasks were delivered 3 times per day and were self-initiated when starting marijuana use. Completion of phone surveys triggered the delivery of cognitive tasks. Participants completed up to 30 days of daily data collection. Multilevel models examined associations between ratings of subjective marijuana high (rated 0-10) and performance on each cognitive task (reaction time and number of correct responses) and tested the number of distractions (rated 0-4) during the mobile task session as a moderator of the association between ratings of subjective marijuana high and task performance.

Results: Participants provided 2703 data points, representing 451 reports (451/2703, 16.7%) of marijuana use. Consistent with slight impairing effects of acute marijuana use, an increase in the average rating of subjective marijuana high was associated with slower average reaction time on all 3 tasks—Flowers (B=2.29; SE 0.86; *P*=.008), marijuana Stroop (B=2.74; SE 1.09; *P*=.01), and DSST (B=3.08; SE 1.41; *P*=.03)—and with fewer correct responses for Flowers (B=−0.03; SE 0.01; *P*=.01) and DSST (B=−0.18; SE 0.07; *P*=.01), but not marijuana Stroop (*P*=.45). Results for distraction as a moderator were statistically significant only for certain cognitive tasks and outcomes. Specifically, as hypothesized, a person's average number of reported distractions moderated the association of the average rating of subjective marijuana high (over and above a session's rating) with the reaction time for marijuana Stroop (B=−52.93; SE 19.38; *P*=.006) and DSST (B=−109.72; SE 42.50; *P*=.01) and the number of correct responses for marijuana Stroop (B=−0.22; SE 0.10; *P*=.02) and DSST (B=4.62; SE 1.81; *P*=.01).

Conclusions: Young adults' performance on mobile cognitive tasks in the natural environment was associated with ratings of acute subjective marijuana high, consistent with slight decreases in cognitive functioning. Monitoring cognitive functioning in real time in the natural environment holds promise for providing immediate feedback to guide personal decision making.

(JMIR Mhealth Uhealth 2020;8(3):e16240) doi:10.2196/16240

KEYWORDS

marijuana; cannabis; cell phone; memory, short-term; cognition

Introduction

Background

Adverse effects of marijuana use on cognitive functioning have been reported by some young adults [1,2], with associated negative consequences such as injury and fatality due to driving while *high* on marijuana [3,4]. The emerging research base on cognitive impairments associated with marijuana use indicates slight and selective cognitive functioning impairments, particularly with certain early onset, heavy, and chronic patterns of marijuana use [5,6]. The effects of marijuana use on cognition, in particular, also depend on factors such as the cognitive domain and whether testing occurs during acute (within 0-6 hours of use) or nonacute (>6 hours since use) periods. Although cross-sectional laboratory studies have compared individuals who use marijuana with healthy controls on measures of cognitive functioning [6], much less is known about the acute effects of marijuana use on cognitive functioning in the natural environment.

Acute Effects of Marijuana on Cognition in Laboratory Studies

For smoked marijuana, the subjective effect of *feeling high* typically begins within 5 min of use and reaches a peak within 30 min, depending on the dose and smoking rate [7]. Laboratory studies indicate that during acute marijuana intoxication, verbal and working memory are typically impaired, and inhibitory control is reduced [8-10]. Findings from these laboratory studies [8-10] on the acute effects of marijuana on cognitive functioning guided the selection of the mobile cognitive tasks used in this study, which assess visuospatial working memory (Flowers task) [11], attentional bias to marijuana-related cues (marijuana Stroop) [12], and information processing and psychomotor speed (digit symbol substitution task [DSST]) [13,14].

Mobile Cognitive Assessment Studies

To date, neuropsychological tests administered in laboratory settings have shown moderate associations with measures of daily functioning [15]. By comparison, data on cognitive functioning collected using ecological momentary assessment (EMA) have greater ecological validity than laboratory assessment as cognitive processes are assessed in real-world contexts [16]. A systematic review of mobile cognitive assessments reported good internal consistency and test-retest reliability for the mobile cognitive assessments studied, in addition to good convergent and divergent validity with laboratory-based measures [17]. The cognitive domain examined most often (7 studies) by mobile assessment was working memory [17]. For example, a 1-week EMA study that administered a mobile visual working memory task multiple (approximately 5-7) times per day to young adult cigarette smokers found that working memory performance decreased with acute marijuana (odds ratio [OR] 0.91, 95% CI 0.84-0.99) and alcohol use (OR 0.87, 95% CI 0.79-0.95) and increased with acute tobacco use (OR 1.11, 95% CI 1.04-1.18) [18]. Although this EMA study provides important insights into acute effects of substance use on working memory in the natural environment, the study focused on young adults who primarily reported cigarette rather than marijuana use; examined only working memory; and included limited information on the level of subjective marijuana high (ie, only examined yes or no reports of use) associated with working memory performance.

Another popular mobile cognitive task uses some version of an addiction Stroop [19]. The addiction Stroop measures attentional bias, the ability of substance-related stimuli to engage attention, particularly among individuals with heavier patterns of substance use [20]. For individuals with cannabis use disorders, cognitive biases for marijuana cues have generally been observed using different methods (eg, visual dot probe, marijuana Stroop) [19]. In a laboratory study examining marijuana Stroop, attentional bias was correlated with both the frequency of marijuana use and subjective craving [12]. Notably, no study to our knowledge has yet reported results for a mobile version of marijuana Stroop. In a mobile version of the alcohol Stroop, attentional bias scores were not associated with individual differences in drinking behavior [21]. In contrast, attentional bias for cigarette smoking cues assessed by a mobile version of the smoking Stroop administered on a personal digital assistant was associated with nicotine dependence severity [22] and nicotine craving during the early stages of a quit attempt [23]. The mixed findings for attentional bias assessed by mobile versions of an addiction Stroop might be due to factors such as the type of substance (alcohol and nicotine) assessed [20] and task parameters (eg, number of trials used in the task).

Two other pilot studies used mobile versions of a Stroop task. One study of outpatients with substance use disorders and healthy controls found practice effects with mobile tasks completed 5 times per day for a week, but only for healthy controls [24]. However, another study that examined a mobile Stroop found no practice effects among participants with methamphetamine dependence or healthy controls who completed mobile tasks twice daily for 2 weeks [25]. Another factor to consider in mobile cognitive assessment is the impact of distraction on task performance [23]. On a cigarette smoking Stroop task, the number of reported interruptions during task performance was associated with slower reaction times and more errors, but interruptions were not associated with the cigarette smoking Stroop effect (ie, slow reaction time when viewing smoking-related words) [23]. These findings suggest

the use of examining distraction as a moderator of mobile task performance, in addition to considering practice effects.

Study Objectives

Informed by laboratory research, this pilot EMA study explored the acute effects of marijuana use on young adults' performance during 3 brief mobile cognitive tasks assessing visuospatial working memory (Flowers task), attentional bias to marijuana-related words (marijuana Stroop), and information processing and psychomotor speed (DSST). Multilevel analyses, conducted separately for each of the 3 cognitive tasks and 2 outcomes (reaction time and number of correct responses), tested the hypothesis that as the rating of momentary subjective marijuana high increases, the reaction time on the mobile cognitive task will slow down and the number of correct responses will decrease. Analyses also examined typical levels of distraction across sessions as a moderator of the association between the typical ratings of subjective marijuana high across sessions and cognitive task performance. Moderation analyses tested the hypothesis that the average number of reported distractions (across sessions) will intensify the effect of being high on marijuana on reaction time and the number of correct responses.

Methods

Recruitment

Young adults (aged 18-25 years) who reported marijuana use at least two times per week (44/71, 62% female) were recruited through a participant registry (Pitt+Me) and Craigslist advertisements in Pittsburgh, Pennsylvania. The exclusion criteria were as follows: currently seeking treatment for substance use, self-reported history of psychosis, and use of medication or a device (eg, pacemaker) that could affect the heart rate.

Participants

Individuals who completed at least five mobile sessions (1 mobile session=1 phone survey + 3 cognitive tasks) were included in the analyses, based on research suggesting that participants gain familiarity with mobile tasks during early sessions (ie, first 5 sessions) [26]. Completion of the phone survey immediately triggered the cognitive tasks. Participants who did not complete at least five sessions (n=3) were excluded. Participants who had scores only when high on marijuana (n=4) were excluded as they do not provide information on session-level comparisons of *high* vs *not high* on marijuana. In addition, 4 participants with missing scores for estimated intellectual functioning (see *Baseline Measures*) were excluded from the analyses. Thus, the analysis sample included 60 participants, of which 37 (62%) were female (mean age 20.0, SD 1.8 years), 45 (75%) were white, 8 (13%) were black, and 7 (12%) were of another race or ethnicity (ie, Asian, Asian Indian, Hispanic, or multiracial). Most participants (40/60, 66%) reported attending some college, 25% (15/60) reported having a high school diploma or equivalent, and 8% (5/60) were college graduates. The majority (55/60, 92%) owned an iOS device, and 8% (5/60) owned an Android mobile phone.

Procedure

Eligible individuals provided written informed consent for study participation. At the baseline assessment, participants installed study mobile apps (eg, AWARE [27] to deliver phone surveys, MUSE [28] to deliver cognitive tasks) on their personal phones, and research staff trained participants on completion of the mobile surveys and cognitive tasks. At baseline, participants completed an interview and questionnaires assessing demographic characteristics, substance use history, and neuropsychological measures assessing attention, memory, and response inhibition. IQ was estimated using a reading test [29]. After baseline, participants completed up to 30 days of daily data collection (see *Compensation*). Daily data collection included scheduled assessments and user-initiated assessments (see *Phone Surveys: Self-Initiated Marijuana Use and Fixed Time Daily Surveys*). At the end of the daily data collection period (phone surveys and mobile cognitive tasks), participants completed a wrap-up session. The University of Pittsburgh institutional review board approved the research protocol.

Baseline Measures

The National Institute on Drug Abuse (NIDA) Quick Screen [30] is a widely used measure to screen 10 types of substance use and substance-related problems covering time frames of lifetime (yes or no) and past 3 months (5-point scale: 0=never to 4=almost daily). Scores of 0 to 3 indicate low risk, 4 to 26 indicate moderate risk, and ≥27 indicate high risk.

The National Adult Reading Test-Revised [29] provided an estimate of full-scale IQ (FSIQ) to account for individual differences in premorbid IQ, which might affect cognitive task performance [10]. A validation study found that National Adult Reading Test FSIQ estimates were similar to Wechsler Adult Intelligence Scale-Revised estimates [31]. The sample mean estimated FSIQ was 110.9 (SD 6.1; range 89.9-121.6).

Phone Surveys: Self-Initiated Marijuana Use and Fixed Time Daily Surveys

Completion of the self-initiated and fixed time daily surveys both immediately triggered the start of mobile cognitive tasks. Participants were instructed to complete self-initiated reports at the start of marijuana use (ie, typically within the first 15 min of initiating use, *when feeling high*). Participants reported the time marijuana use started, mode of use (eg, joint, vape, pen, and bowl), quantity consumed (eg, grams or hits), on the question "How high are you feeling right now?" (0=none to 10=a lot), and other substance use (eg, number of drinks consumed).

Fixed-time daily surveys were delivered 3 times per day (ie, 10 AM, 3 PM, and 8 PM) with a 5-hour window for completion. Participants received a notification that the survey and tasks were available but did not receive reminders to complete the survey. Fixed-time surveys (similar to self-initiated reports) included items on time of last marijuana use, quantity consumed, the question "How high are you feeling right now?" and other items (eg, mood rating). Survey completion immediately triggered the administration of the 3 cognitive tasks in a randomized order. A session (a phone survey and mobile cognitive tasks) *timed out* if there was a lag in response for >1

min, which would end the session, such that remaining tasks and post-task survey items (eg, distraction item, see below) could not be done. With this schedule of fixed time and self-initiated assessments, participants reported their rating of subjective marijuana high immediately before performing the mobile cognitive tasks, which permitted the examination of task performance when participants reported not being high (subjective high rating=0) relative to reports when feeling *high* (subjective high rating>0).

Mobile Phone Cognitive Tasks and Rating of Distraction After Session Completion

The 3 brief cognitive tasks (approximately 5 min in total to complete) included the following: visuospatial working memory task (Flowers task [11]), marijuana Stroop [12], and DSST [13]. The Flowers task and marijuana Stroop provided immediate feedback on incorrect responses. The DSST did not provide any feedback on correct or incorrect responses to minimize distractions during task performance. The 3 tasks did not provide a score regarding performance.

The Flowers task assesses short-term visuospatial working memory [11]. Participants watched flowers in a grid light up one at a time and were instructed to replicate the sequence by touching the flowers in the grid in the same order (Figure 1). The task adapts to a test taker's ability and increases or decreases its difficulty, starting with a 3×3 grid and increasing to a 4×4 grid with success or decreasing in difficulty with error. The task ends with 2 consecutive errors or a maximum of 6 correct responses. Previous work found that task performance (ie, number of correct responses) distinguished patients with Parkinson disease from healthy controls [26]. The Flowers task scores are the number of correct responses [11] and reaction time. Reaction time was added to assess possible psychomotor slowing associated with acute marijuana use, similar to the other 2 mobile tasks. Embedded sensors in mobile phones allow precise measurement for reaction time tasks [32].

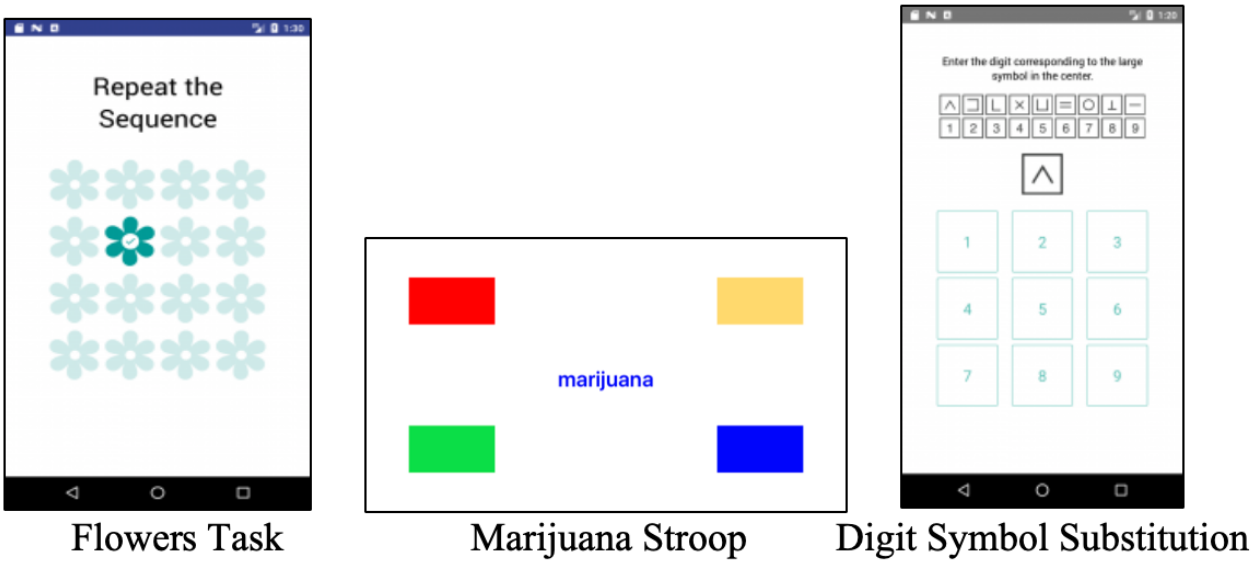
Marijuana Stroop [12] measures attentional bias for marijuana-related stimuli. Participants were presented with 14 marijuana-related words (eg, hash and joint) and 14 neutral words (eg, sand and winds) in a randomized order in 2 blocks. Words appeared in 4 colors (red, yellow, blue, and green). Word color was random, but each color was shown at least once in each set of 28 words. Participants tapped the color of the word as fast as possible (Figure 1). Errors were shown by a red X immediately after the response. A computerized version of the marijuana Stroop task indicated that marijuana-related words captured the attention of marijuana-dependent individuals (ie, longer reaction time for marijuana-related words vs neutral words), but not healthy controls [12]. A mobile alcohol Stroop task had acceptable internal consistency reliability in real-world settings (Cronbach alpha=.70 to .74), with participants showing attentional bias to alcohol words [21].

To compute the marijuana Stroop reaction time score, trials with incorrect responses were excluded, and trials with unrealistically fast (<100 ms; none excluded) or slow (>1250 ms) reaction times were excluded (88/1467, 6.00% of total responses) [33,34]. Internal consistency reliability (Cronbach alpha) for response latency was computed by averaging within each trial type [34] for marijuana-related words and neutral words (alpha=.92 and .92, respectively). The total number of correct responses in the marijuana Stroop task also excluded trials with unrealistically fast or slow reaction times. There were no significant results from multilevel analyses for a marijuana Stroop effect (results not shown), computed as the difference in reaction time for marijuana-related and neutral words [12]. Preliminary multilevel analyses, which examined marijuana Stroop reaction time scores for combined marijuana and neutral words and separately by word type (marijuana-related and neutral) [12], indicated similar results for combined and separate word types. Similarly, preliminary multilevel results for the number of correct responses were similar for a combination of marijuana-related and neutral words and separate word types. Thus, marijuana Stroop scores are reaction time and the total number of correct responses (a combination of marijuana-related and neutral words).

DSST [13] measures information processing and psychomotor processing speed and is sensitive to acute drug effects [14]. The DSST requires quick response to visual symbols by touching the corresponding digit (1-9) shown in the reference key (Figure 1). New reference keys provided after each response minimize learning effects within the 60-second task session. No feedback was provided regarding a correct or an incorrect response. DSST scores are reaction time and the number of correct responses.

The number of distractions while performing the cognitive tasks was reported after completing the 3 tasks by responding to the following item: "How many times were you distracted during completion of the tasks?" (coded 0 to 4 or more times; 0-4) [23].

Figure 1. Screenshots of Flowers task, marijuana Stroop, and digit symbol substitution task.



Compensation

Participants were compensated US $75 for completing the baseline assessment. During the first 14 days of daily data collection, for each day on which >75% of data were collected (eg, phone surveys and cognitive tasks), participants earned US $10; if <75% of data were collected on a particular day, no money was earned. If the participant had good compliance during the first 14 days of data collection and was willing to continue for another 14 days, data collection continued for a second 14-day period at the same compensation rate. Participants earned US $25 for the wrap-up session (user experience interview and final data download).

Statistical Analysis

To examine the acute effects of marijuana use on mobile cognitive task performance, generalized linear mixed effects models were fit to the data (xtmixed: Stata Statistical Software 15.0, StataCorp LLC) [35]. This multilevel modeling approach can accommodate mixed (fixed and random) effects across multiple data levels and account for the correlations between repeated measures (ie, sessions: level 1) within participants (level 2). Mixed models used maximum likelihood estimation, leveraging all available data to accommodate missing data. Likelihood ratio testing (−2LL difference between models) evaluated the statistical significance of nested models when random effects were added, and Akaike and Bayesian Information Criteria were used to evaluate the model fit between non-nested models [36].

The main outcomes of reaction time and number of correct responses were examined for each of the 3 mobile cognitive tasks. Distributions for the outcome variables of the number of correct responses for the Flowers and marijuana Stroop tasks indicated negative skew. Exponential (cubic) transformation reduced skew but did not normalize distributions and resulted in similar findings. Thus, untransformed results are reported. Analyses used the first 60 completed sessions (sparse data at >60 sessions). Alpha was set at $P<.05$ without protecting the family-wise alpha rate for multiplicity in this pilot study.

Preliminary analyses examined correlations, computed intraclass correlations (ICCs), and modeled time trends (ie, practice effects; see Figures 2 and 3) for the outcomes (reaction time and number of correct responses). ICCs and time trends were examined for time-varying predictors (eg, subjective marijuana high and distraction) in unconditional models. As experience with the cognitive task itself (rather than the passage of time) likely contributes to a *practice* effect, the sequential count of completed sessions was used as the measure of time [36]. Time was coded so that the first completed session (done at baseline) represented session=0 in all models. For both outcomes, the model that provided the best fit was a linear model for Flowers and marijuana Stroop tasks and a quadratic model for DSST.

In total, 2 time-varying predictors were examined: subjective marijuana high (“How high are you feeling right now?” rated 0-10) and distraction (“How many times were you distracted during completion of the tasks?” coded 0-4) and their interaction. Decomposition of person-level (level 2) and session-level (level 1) effects in the mixed model was done as follows: the time-varying covariates (eg, subjective marijuana high and distraction) were centered at a constant (ie, 0; *constant-centered [CC]*) and tested in the model as level 1 predictors; and the corresponding person mean counterpart was entered simultaneously at level 2 [36]. When the CC and person mean–centered variables are entered together in the model, the CC variable represents session-to-session variation. The person mean variable represents the unique effect of the person’s average level of that variable on the outcome over and above the *absolute amount* of the time-varying CC effect, or in other words, individual differences in subjective marijuana high or distraction across all sessions [36].

Subjective marijuana high was centered such that a rating of 5=*moderately high* was recoded to be centered at 0 (new range −5 to 5). Ratings of subjective high had an ICC of 0.04 and showed no systematic change over time. Ratings for distraction were not transformed (range 0-4; ICC=0.25) and were *centered* at 0, and also showed no systematic change over time. Person means for subjective marijuana high (referred to as *subjective*

high person mean) and distraction (*distraction person mean*) were computed.

Static covariates entered at the person level (level 2) included gender (0=male and 1=female), age (0=age 20), and FSIQ (0=estimated FSIQ of 110). Interactions of time-varying predictors (eg, subjective marijuana high, distraction) with static covariates were not tested to limit multiple comparisons and because no *a priori* hypotheses for these interactions were proposed. Other drug use was explored for inclusion as a time-varying predictor but was highly collinear with a rating of subjective marijuana high and was not included (334/451, 74.0% of phone surveys reported no alcohol use; 408/451, 90.5% reported no nicotine use). Weekend or weekday use (see Multimedia Appendix 1) was not a significant predictor.

Likelihood ratio testing, which examined relative model fit when including random slopes for time and time-varying covariates (eg, subjective marijuana high and distraction), indicated that including random slopes for a session and subjective marijuana high did not improve the model fit. A random slope for distraction for the outcome of reaction time improved the model fit (see Multimedia Appendix 1) and was included in models for this outcome.

Figure 2. Average reaction time across sessions: flowers, marijuana Stroop, and digit symbol substitution task tasks. DSST: digit symbol substitution task.

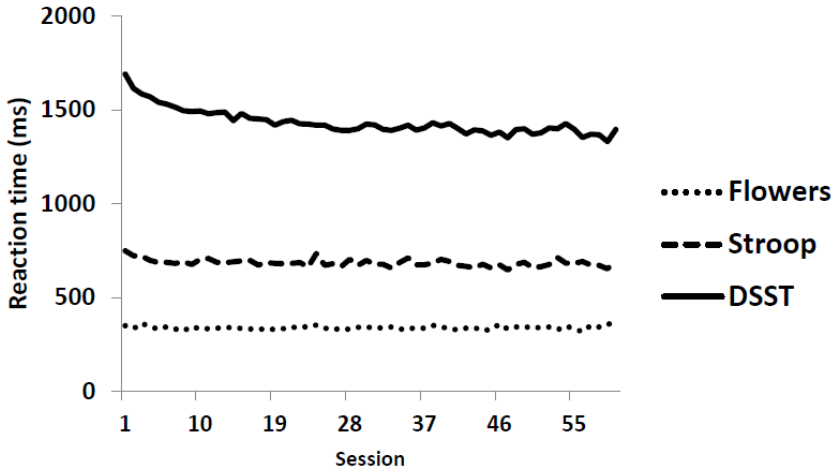
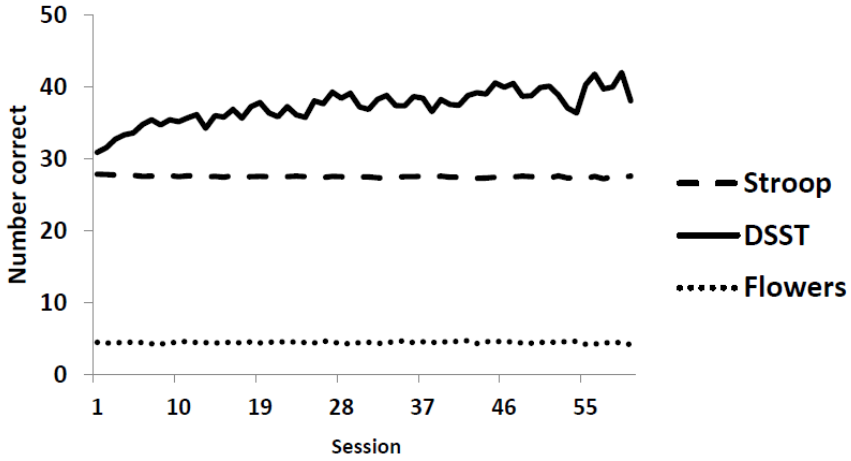


Figure 3. Average number of correct responses across sessions: flowers, marijuana Stroop, and digit symbol substitution task tasks. DSST: digit symbol substitution task.



Results

Marijuana Use in the Sample

In the analysis sample, the average age of onset for marijuana use was 16.5 years (SD 1.7; range 13-22), and the average age of onset for regular marijuana use (ie, using marijuana at least once per month for at least six months) was 17.2 years (SD 1.6). At baseline, 27% (16/60) of participants reported daily marijuana use, 10% (6/60) reported *almost daily* use (5-6 times per week), 33% (20/60) reported use 3 to 4 times per week, and 30% (18/60) reported use 2 times per week. The mean score on the NIDA Quick Screen was 15.5 (SD 6.4) [30], indicating moderate risk associated with marijuana use. Almost all (54/57, 95%) participants scored in the moderate risk range, with the remaining (6/120, 5%) participants scoring in the severe range.

Across 60 sessions completed over 5 to 30 days (mean 20.6 days, SD 6.5; Multimedia Appendix 2), 2703 data points were obtained from 60 young adults who provided 451 reports (451/2703, 16.7%) of marijuana use. Reports were obtained from morning (133/451, 29.4%), afternoon (115/451, 25.6%),

evening (106/451, 23.4%), and self-initiated marijuana surveys (97/451, 21.6%). The average number of sessions completed per participant was 25.3 (SD 16.2; Multimedia Appendix 3). Participants completed 52.98% (2119/4000) of fixed time assessments.

The average number of sessions associated with marijuana use was 7.0 (SD 5.8). During sessions completed when high on marijuana (*any* rating of feeling *high*; n=451), the average level of subjective marijuana high was rated 4.7 out of 10 (SD 2.2; range 1-10). The most common method of consumption was pen or vaporizer (207/451, 45.9%), followed by bong (100/451, 22.2%), bowl or pipe (59/451, 13.0%), joint (38/451, 8.5%), blunt (38/451, 8.4%), edible (5/451, 1.1%), or tincture (4/451, 0.8%). The average quantity consumed per occasion was 0.8 g (SD 1.1), and when reported as hits, an average of 6.0 hits (SD 11.6). There was a small positive correlation (r=0.13; P=.001) between the number of hits reported and the rating of subjective marijuana high, but no statistically significant association between the number of grams reported and rating of subjective marijuana high (r=−0.03; P=.58).

The average number of distractions reported during a mobile cognitive task session when not high was 1.1 (SD 1.1; range 0-4), and when high was 0.9 distractions (SD 1.0; range 0-4). There was a very small negative correlation (r=−0.04; P=.046) between the number of distractions reported and the rating of subjective marijuana high.

Mobile Cognitive Tasks

Tables 1 and 2 show the intercorrelations among the mobile tasks for reaction time and number of correct responses, based on a subjective marijuana high rating of 0 (*not high*) vs a rating from 1 to 10 (when feeling *high*). Reaction times were positively correlated for all 3 tasks (r=0.18 to 0.48; P=.001). The number of correct responses was positively correlated for Flowers and DSST (r=0.26 to 0.27; P=.001), negatively correlated for DSST and marijuana Stroop (r=−0.14 to −0.18; P=.001) and showed no association for marijuana Stroop and Flowers task (P>.36). Raw differences in reaction times when high vs not high on the tasks were small (<50 ms), as were differences in the number of correct responses, generally indicating slightly slower reaction times and slightly fewer correct responses when high on marijuana (vs not high).

For reaction time, when not high on marijuana (subjective high rating=0), ICCs were as follows: DSST=0.57, marijuana Stroop=0.23, and Flowers=0.38. When high on marijuana (subjective high rating>0), ICCs for reaction time were as follows: DSST=0.61, marijuana Stroop=0.25, and Flowers=0.29. The ICCs indicated that between 23% and 61% of the variances in reaction time for mobile tasks were between persons.

For number of correct responses, when not high on marijuana, ICCs were as follows: DSST=0.49, marijuana Stroop=0.13, and Flowers=0.15. When high on marijuana, ICCs for number of correct responses were as follows: DSST=0.50, marijuana Stroop=0.11, and Flowers=0.27. For number of correct responses, 11% to 50% of the variances in mobile task performance were between persons. The generally lower ICCs for marijuana Stroop and Flowers tasks (ie, reaction time and number of correct responses) suggested that their overall variance primarily reflects within-person, session-to-session fluctuations, providing the rationale for a multilevel analysis.

Table 1. Pearson correlations (r) and mean reaction time (millisecond) for 3 cognitive tasks (correlations do not take clustering of cases within individuals over time into account).

Measure	Not high (subjective high rating=0), n=2159 sessions		High (subjective high rating >0), n=389 sessions		Not high	High
	Flowers	Marijuana Stroop	Flowers	Marijuana Stroop	DSST[a]	DSST
Marijuana Stroop	0.20[b]	N/A[c]	0.34 [b]	N/A	N/A	N/A
DSST	0.18[b]	0.36[b]	0.25[b]	0.48[b]	N/A	N/A
Reaction time, mean (SD)	335.80 (73.38)	682.08 (115.24)	354.87 (80.51)	713.60 (139.73)	1440.83 (170.59)	1488.75 (187.56)

[a]DSST: digit symbol substitution task.

[b]P<.001.

[c]Not applicable.

Table 2. Pearson correlations (r) and mean number of correct responses for 3 cognitive tasks (correlations do not take clustering of cases within individuals over time into account).

Measure	Not high (subjective high rating=0), n=2252 sessions		High (subjective high rating >0), n=451 sessions		Not high	High
	Flowers	Marijuana Stroop	Flowers	Marijuana Stroop	DSST[a]	DSST
Marijuana Stroop	0.02	N/A[b]	−0.02	N/A	N/A	N/A
DSST	0.26[c]	−0.14[c]	0.27[c]	−0.18[c]	N/A	N/A
Correct responses						
Mean (SD)	4.49 (0.88)	27.52 (0.77)	4.31 (1.09)	27.51 (0.81)	37.03 (7.76)	35.44 (8.58)
Median (SD)	5.00 (1.00)	28.00 (1.00)	5.00 (1.00)	28.00 (1.00)	38.00 (9.00)	37.00 (10.00)

[a]DSST: digit symbol substitution task.

[b]Not applicable.

[c]*P*<.001.

Reaction Time: Associations With Subjective Marijuana High and Distraction

For all 3 mobile cognitive tasks, there was a significant session-level association of subjective marijuana high with reaction time (Tables 3-5), such that an increase in the rating of subjective marijuana high was associated with slower reaction time: Flowers task (B=2.29; SE 0.86; *P*=.008), marijuana Stroop (B=2.74; SE 1.09; *P*=.01), and DSST (B=3.08; SE 1.41; *P*=.03). In addition, there was a significant effect of subjective high person mean for both marijuana Stroop (B=77.78; SE 25.48; *P*=.002), and DSST (B=181.32; SE 55.83; *P*=.001), indicating an effect over and above that of a specific session for slower (ie, increasing) reaction time with a greater average rating of subjective marijuana high across sessions (ie, the person's *usual* level of high, relative to other people with lower ratings).

The sample size of the Flowers task has 2 fewer cases compared with the other 2 tasks because of the late initiation of reaction time data collection due to programming delay.

For both marijuana Stroop and DSST, there was a significant interaction of subjective marijuana high person mean with distraction person mean: marijuana Stroop (B=−52.92; SE 19.38; *P*=.006) and DSST (B=−109.72; SE 42.50; *P*=.01; Tables 4 and 5). For both tasks (Figures 4 and 5), there was little estimated difference in reaction time at low average levels of subjective marijuana high person mean, but contrary to the hypothesis, at higher average levels of distraction person mean and higher average ratings of subjective marijuana high person mean, reaction time was estimated to decrease (with wide 95% CIs at the highest average level of contextual distraction). In contrast, at low average levels of distraction person mean, reaction time was predicted to be slower as the subjective rating of high person mean increased (possibly reflecting the effect of marijuana on the slowing of psychomotor functioning).

Table 3. Flowers task: multilevel model of marijuana high in relation to reaction time.

Effects	Flowers reaction time (n=58)			
	Estimate	SE	95% CI	*P* value
Fixed effects				
Person level (level 2)				
Intercept	592.88	128.62	N/A[a]	.001[b]
Session	−0.14	0.08	N/A	.08
Session^2 (quadratic)	N/A	N/A	N/A	N/A
Subjective high (PM[c])	57.45	30.16	N/A	.06
Distraction (PM)	−81.80	89.75	N/A	.36
Subjective high (PM) × distraction (PM)	−20.74	20.79	N/A	.32
Gender (0=male, 1=female)	−9.64	10.59	N/A	.36
Age (0=age 20)	6.49	3.17	N/A	.04[d]
Full-scale IQ (0=IQ score of 110)	2.50	1.04	N/A	.02[d]
Session level (level 1)				
Subjective high (CC[e])	2.29	0.86	N/A	.008[b]
Distraction (CC)	−4.38	2.88	N/A	.13
Subjective high (CC) × distraction (CC)	−0.39	0.55	N/A	.48
Random effects				
Level 1 residual variance	3415.85	99.90	3225.55 to 3617.38	N/A
Intercept	1259.15	274.92	820.79 to 1931.64	N/A
Distraction	56.58	28.40	21.15 to 151.33	N/A
Covariance (distraction, intercept)	0.48	71.70	−140.04 to 141.00	N/A

[a]Not applicable.

[b]*P*<.01.

[c]PM: person mean scores, reflecting individual differences in subjective marijuana high or distraction across all sessions.

[d]*P*<.05.

[e]CC: constant-centered scores (centered at 0), reflecting session-to-session variation in scores or session-specific scores.

Table 4. Marijuana Stroop: multilevel model of marijuana high in relation to reaction time.

Effects	Stroop reaction time (n=60)			
	Estimate	SE	95% CI	*P* value
Fixed effects				
Person level (level 2)				
Intercept	1018.27	106.73	N/A[a]	.001[b]
Session	−0.61	0.10	N/A	.001[b]
Session^2 (quadratic)	N/A	N/A	N/A	N/A
Subjective high (PM[c])	77.78	25.48	N/A	.002[b]
Distraction (PM)	−211.10	82.07	N/A	.01[d]
Subjective high (PM) × distraction (PM)	−52.92	19.38	N/A	.006[b]
Gender (0=male, 1=female)	2.40	14.31	N/A	.87
Age (0=age 20)	12.78	4.08	N/A	.002[b]
Full-scale IQ (0=IQ score of 110)	−0.03	1.38	N/A	.98
Session level (level 1)				
Subjective high (CC[e])	2.74	1.09	N/A	.01[d]
Distraction (CC)	−0.44	3.94	N/A	.91
Subjective high (CC) × distraction (CC)	−10.32	0.72	N/A	.07
Random effects				
Level 1 residual variance	5858.83	168.35	5537.99 to 6198.26	N/A
Intercept	2647.94	574.67	1730.52 to 4051.72	N/A
Distraction	188.00	67.35	93.16 to 379.40	N/A
Covariance (distraction, intercept)	−162.67	158.49	−473.31 to 147.96	N/A

[a]Not applicable.

[b]*P*<.01.

[c]PM: person mean scores, reflecting individual differences in subjective marijuana high or distraction across all sessions.

[d]*P*<.05.

[e]CC: constant-centered scores (centered at 0), reflecting session-to-session variation in scores or session-specific scores.

Table 5. Digit symbol substitution task: Multilevel model of marijuana high in relation to reaction time.

Effects	DSST[a] reaction time (n=60)			
	Estimate	SE	95% CI	*P* value
Fixed effects				
Person level (level 2)				
Intercept	2328.14	233.72	N/A[b]	.001[c]
Session	−9.45	0.45	N/A	.001[c]
Session^2 (quadratic)	0.12	0.01	N/A	.001[c]
Subjective high (PM[d])	181.32	55.83	N/A	.001[c]
Distraction (PM)	−441.54	179.82	N/A	.01[e]
Subjective high (PM) × distraction (PM)	−109.72	42.5	N/A	.01[e]
Gender (0=male, 1=female)	−2.75	31.61	N/A	.93
Age (0=age 20)	7.55	8.92	N/A	.4
Full-scale IQ (0=IQ score of 110)	4.91	3.02	N/A	.1
Session level (level 1)				
Subjective high (CC[f])	3.08	1.41	N/A	.03[e]
Distraction (CC)	2.21	4.6	N/A	.63
Subjective high (CC) × distraction (CC)	−1.01	0.92	N/A	.27
Random effects				
Level 1 residual variance	9950.65	285.83	9405.91 to 10,526.95	N/A
Intercept	12906.21	2506.7	8820.12 to 18,885.27	N/A
Distraction	63.55	59.06	10.28 to 392.83	N/A
Covariance (distraction, intercept)	−67.68	330.49	−715.43 to 580.07	N/A

[a]DSST: digit symbol substitution task.

[b]Not applicable.

[c]*P*<.01.

[d]PM: person mean scores, reflecting individual differences in subjective marijuana high or distraction across all sessions.

[e]*P*<.05.

[f]CC: constant-centered scores (centered at 0), reflecting session-to-session variation in scores or session-specific scores.

Figure 4. Marijuana Stroop: distraction person mean (PM) as a moderator of the association between subjective high PM and reaction time. Distraction PM at low (0), moderate (2) and high levels (4). PM: person mean.

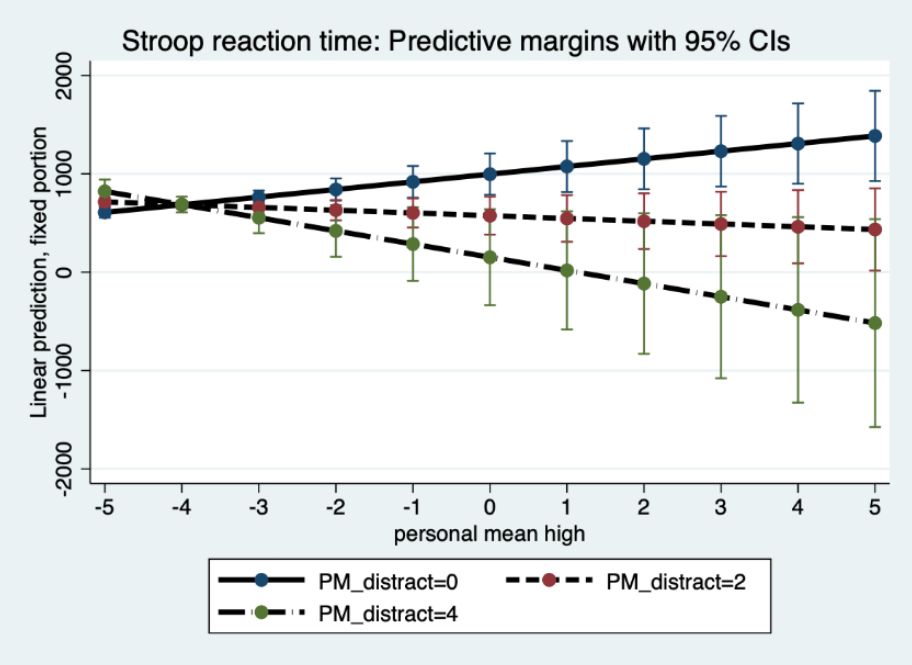
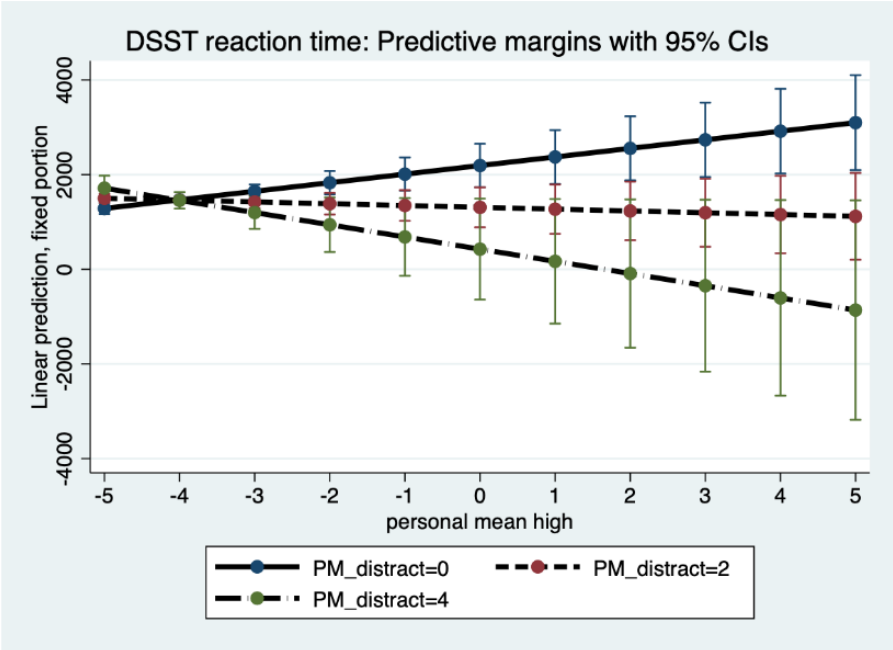


Figure 5. Digit symbol substitution task: distraction person mean (PM) as a moderator of the association between subjective high PM and reaction time. Distraction PM at low (0), moderate (2) and high levels (4). DSST: digit symbol substitution task; PM: person mean.



Number of Correct Responses: Associations With Subjective Marijuana High and Distraction

There was a significant session-level association of subjective marijuana high with the number of correct responses for Flowers (B=−0.03; SE 0.01; *P*=.01) and DSST (B=−0.18; SE 0.07; *P*=.01) tasks, such that the increase in the rating of subjective marijuana high was associated with fewer correct responses (eg, Table 6). Although the session-level effect of subjective marijuana high on number of correct responses for marijuana Stroop was not significant (*P*=.45; Table 7), the effect of subjective marijuana high person mean was statistically significant (B=0.37; SE 0.13; *P*=.003), indicating unique effects of the person's average subjective rating of marijuana high across sessions controlling for a specific occasion's rating of subjective marijuana high on the number of correct responses in the marijuana Stroop task. For DSST (Table 8), the effect of subjective marijuana high person mean was also significant (B=−6.83; SD 2.38, *P*=.004), indicating that at an average subjective rating of marijuana high, the number of correct DSST responses was lower (controlling for a specific occasion's rating of subjective marijuana high).

Table 6. Flowers task: multilevel model of marijuana high in relation to number of correct responses.

Effects	Flower number of correct responses (n=58)			
	Estimate	SE	95% CI	*P* value
Fixed effects				
Person level (level 2)				
Intercept	2.92	0.85	N/A[a]	.001[b]
Session	0	0	N/A	.007[b]
Session^2 (quadratic)	N/A	N/A	N/A	N/A
Subjective high (PM[c])	−0.35	0.2	N/A	.08
Distraction (PM)	0.83	0.65	N/A	.2
Subjective high (PM) × distraction (PM)	0.18	0.15	N/A	.23
Gender (0=male, 1=female)	−0.05	0.11	N/A	.64
Age (0=age 20)	−0.14	0.03	N/A	.001[b]
Full-scale IQ (0=IQ score of 110)	0	0.01	N/A	.66
Session level (level 1)				
Subjective high (CC[d])	−0.03	0.01	N/A	.01[e]
Distraction (CC)	−0.13	0.03	N/A	.001[b]
Subjective high (CC) × distraction (CC)	0.01	0.01	N/A	.21
Random effects				
Level 1 residual variance	0.59	0.02	0.56 to 0.63	N/A
Intercept	0.16	0.03	0.10 to 0.24	N/A

[a]Not applicable.

[b]*P*<.01.

[c]PM: person mean scores, reflecting individual differences in subjective marijuana high or distraction across all sessions.

[d]CC: constant-centered scores (centered at 0), reflecting session-to-session variation in scores or session-specific scores.

[e]*P*<.05.

Table 7. Marijuana Stroop task: multilevel model of marijuana high in relation to number of correct responses.

Effects	Stroop number of correct responses (n=60)			
	Estimate	SE	95% CI	*P* value
Fixed effects				
Person level (level 2)				
Intercept	29.03	0.53	N/A[a]	.001[b]
Session	−0.00	0	N/A	.001[b]
Session^2 (quadratic)	N/A	N/A	N/A	N/A
Subjective high (PM[c])	0.37	0.13	N/A	.003[b]
Distraction (PM)	−0.82	0.41	N/A	.046[d]
Subjective high (PM) × distraction (PM)	−0.22	0.1	N/A	.02[d]
Gender (0=male, 1=female)	0.11	0.07	N/A	.14
Age (0=age 20)	−0.03	0.02	N/A	.11
Full-scale IQ (0=IQ score of 110)	0.02	0.01	N/A	.02[d]
Session level (level 1)				
Subjective high (CC[e])	−0.01	0.01	N/A	.45
Distraction (CC)	−0.10	0.03	N/A	.001[b]
Subjective high (CC) × distraction (CC)	−0.01	0.01	N/A	.23
Random effects				
Level 1 residual variance	0.51	0.01	0.49 to 0.54	N/A
Intercept	0.06	0.01	0.04 to 0.09	N/A

[a]Not applicable.

[b]*P*<.01.

[c]PM: person mean scores, reflecting individual differences in subjective marijuana high or distraction across all sessions.

[d]*P*<.05.

[e]CC: constant-centered scores (centered at 0), reflecting session-to-session variation in scores or session-specific scores.

Table 8. Digit symbol substitution task: multilevel model of marijuana high in relation to number of correct responses.

Effects	DSST[a] number of correct responses (n=60)			
	Estimate	SE	95% CI	*P* value
Fixed effects				
Person level (level 2)				
Intercept	5.81	9.95	N/A[b]	.56
Session	0.25	0.02	N/A	.001[c]
Session^2 (quadratic)	−0.00	0	N/A	.001[c]
Subjective high (PM[d])	−6.83	2.38	N/A	.004[c]
Distraction (PM)	18.78	7.66	N/A	.01[e]
Subjective high (PM) × distraction (PM)	4.62	1.81	N/A	.01[e]
Gender (0=male, 1=female)	−0.23	1.34	N/A	.86
Age (0=age 20)	−0.40	0.38	N/A	.29
Full-scale IQ (0=IQ score of 110)	−0.09	0.13	N/A	.47
Session level (level 1)				
Subjective high (CC[f])	−0.18	0.07	N/A	.01[e]
Distraction (CC)	−0.86	0.22	N/A	.001[c]
Subjective high (CC) × distraction (CC)	0.08	0.04	N/A	.07
Random effects				
Level 1 residual variance	24.66	0.7	23.33 to 26.08	N/A
Intercept	23.21	4.37	16.04 to 33.57	N/A

[a]DSST: digit symbol substitution task.

[b]Not applicable.

[c]*P*<.01.

[d]PM: person mean scores, reflecting individual differences in subjective marijuana high or distraction across all sessions.

[e]*P*<.05.

[f]CC: constant-centered scores (centered at 0), reflecting session-to-session variation in scores or session-specific scores.

For all 3 tasks, the effect of increasing session-level distraction was significantly associated with fewer correct responses: Flowers task (B=−0.13; SE 0.03; *P*=.001), marijuana Stroop (B=−0.10; SE 0.03; *P*=.001), and DSST (B=−0.86; SE 0.22; *P*=.001). The association of distraction person mean with the number of correct responses was significant for marijuana Stroop (B=−0.82; SE 0.41; *P*=.046) and DSST (B=18.78; SE 7.66; *P*=.01) but in opposite directions. Specifically, for marijuana Stroop, increasing distraction person mean had a unique association with fewer correct responses (controlling for a specific occasion's rating of distraction), whereas for DSST, increasing distraction person mean predicted an increase in the number of correct DSST responses (over and above a given session's rating of distraction).

As found for reaction time, in number of correct responses for both marijuana Stroop and DSST, there was a significant interaction of subjective marijuana high person mean with distraction person mean: marijuana Stroop (B=−0.22; SE 0.10; *P*=.02) and DSST (B=4.62; SE 1.81; *P*=.01). For both tasks (Figures 6 and 7), when the subjective marijuana high person mean was on average low, number of correct responses was estimated to be similar across levels of the distraction person mean. When the average level of distraction person mean was high, however, number of correct responses for the marijuana Stroop was estimated to decrease with increasing subjective high person mean. In contrast, for DSST and contrary to the hypothesis, when the average level of distraction person mean was high, number of correct responses for DSST was estimated to show some increase as the average subjective marijuana high person mean increased (although the 95% CI was wide at the highest levels of contextual distraction).

Figure 6. Marijuana Stroop: distraction person mean (PM) as a moderator of the association between subjective high PM and number of correct responses. Distraction PM at low (0), moderate (2) and high levels (4). PM: person mean.

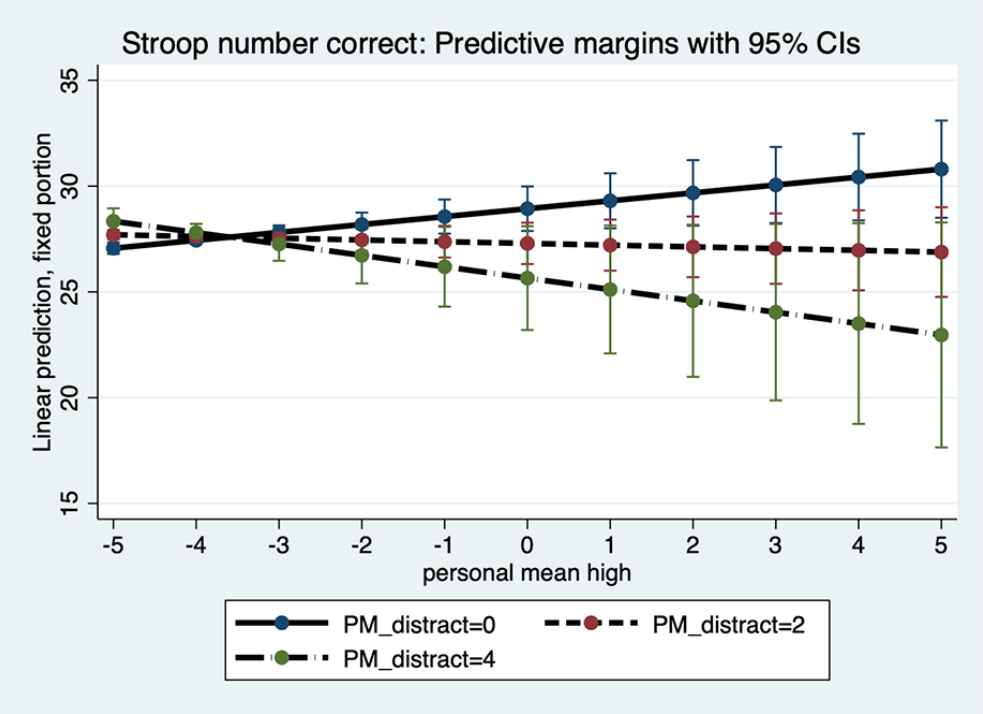
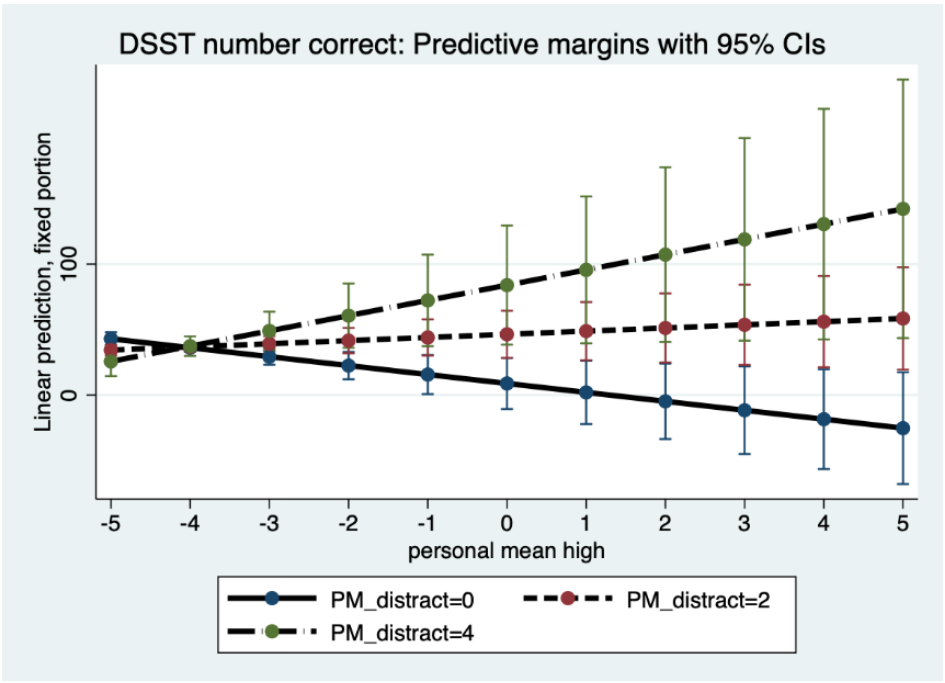


Figure 7. Digit symbol substitution task: distraction person mean (PM) as a moderator of the association between subjective high PM and number of correct responses. Distraction PM at low (0), moderate (2), and high levels (4). DSST: digit symbol substitution task; PM: person mean.



Discussion

Principal Findings

Young adults in this pilot study reported, on average, a *moderate* level of subjective marijuana high when using marijuana (mean 4.7; range 1-10), and roughly one distraction, on average, when completing brief mobile cognitive tasks in the natural environment. For all 3 mobile cognitive tasks, as the average rating of subjective marijuana high increased, average reaction time showed a statistically significant increase, suggesting that the mobile tasks were sensitive to psychomotor slowing associated with acute marijuana use in the natural environment, above and beyond practice effects. The statistically significant acute effects of marijuana in relation to reaction time were small, in the context of average *moderate* ratings of subjective marijuana high. Furthermore, for the Flowers task and DSST, individuals with a greater average rating of subjective marijuana high had, at a statistically significant level, slightly fewer correct responses, compared with those with lower average subjective

marijuana high ratings, further suggesting sensitivity of these 2 mobile tasks to acute marijuana effects collected *in the wild*, as well as some individual differences in marijuana effects. In the uncontrolled daily life testing situations in which the mobile tasks were completed, distraction ratings were uniquely, significantly associated with only certain aspects of task performance (eg, number of correct responses, controlling for covariates) and also significantly moderated the association between ratings of subjective marijuana high and task performance, albeit in ways that were sometimes contrary to prediction, and depended on specific task characteristics in this pilot study.

A consistent finding across the mobile tasks was that, on average, an increasing rating of subjective marijuana high was significantly associated with slower average reaction time. Although other studies have shown slower reaction time for mobile cigarette smoking Stroop [22,23], this is the first study, to our knowledge, to show significantly slower reaction time for a mobile marijuana Stroop in relation to a numerically rated measure (0-10 scale) of subjective marijuana high. Importantly, ratings of subjective marijuana high showed a small, significant correlation with the number of hits, providing some support for the validity of this study's subjective marijuana high measure. However, there was no correlation between reports of grams consumed and subjective marijuana high. The absence of a correlation between the quantity reported in grams and the rating of subjective marijuana high suggests individual differences in tolerance to marijuana. Notably, the assessment of self-reported marijuana quantity is challenging [37] and warrants further study, with previous mobile cognitive assessment limited to only reporting any marijuana use (yes or no) [18].

Results also indicated, for the marijuana Stroop task and DSST, that greater the average rating of subjective marijuana high, slower the response time, at a statistically significant level, after controlling for that session's rating of subjective marijuana high. Thus, a person's typical level of marijuana use appears to have an effect on response time to these mobile cognitive tasks over and above ratings of acute subjective marijuana high at each session, suggesting possible unique effects of, for example, a person's pattern of chronic marijuana use on a specific indicator of task performance [6].

Variations in the test environment, such as distractions, can influence task performance [26]. The slightly lower average number of distractions reported when high (vs not high), and very small negative correlation (r=−0.04; P=.046) between the number of distractions and subjective marijuana high ratings, might reflect that some individuals use marijuana specifically to take a break (eg, *relax*) from distractions in the environment. Alternatively, acute effects of marijuana might reduce the awareness of peripheral distractions for some individuals in certain contexts. The effect of a person's average level of distraction on number of correct responses was significant for the marijuana Stroop task and DSST, but in opposite directions. This finding suggests the importance of considering how task demand characteristics, such as task complexity, and other factors (eg, motivation, effort), including improved measurement of distraction (ie, multi-item self-report and objective measure), are associated with mobile task performance.

Significant interactions of subjective marijuana high and distraction were found only for the marijuana Stroop task and DSST, for both reaction time and number of correct responses, in this pilot study. For the reaction time outcome, both the marijuana Stroop task and DSST showed a similar pattern of results for the distraction interaction. Specifically, as predicted, and consistent with acute effects of marijuana use on slowing of psychomotor functioning, a person who reported a low average level of distraction had slower reaction time as the average subjective marijuana high rating increased. However, contrary to prediction a person who reported a high average level of distraction was estimated (with wide CIs, suggesting a cautious interpretation) to have *faster* reaction time on these 2 tasks as the average subjective marijuana high rating increased. Faster reaction time suggests the possibility of impulsive responding for some individuals on certain tasks. Regarding the outcome of number of correct responses, results of the distraction interaction differed for the marijuana Stroop task and DSST, and only provided partial support for the hypotheses. Given that the results for some distraction interaction hypotheses were contrary to prediction, interpretation of the distraction interaction results warrants caution due to wide 95% CIs at the highest levels of distraction. The mixed findings for the distraction interactions highlight the need for improved distraction measurement, given self-reporting using a single distraction item.

Limitations

This pilot study had limitations. On average, this young adult sample was well educated and reported a moderate level of marijuana-related risk, limiting generalizability. Compliance could be improved (eg, no reminders given), and technical issues (eg, app crashes) reduced completion rates. The 5-hour window to complete fixed time assessments accommodated individual schedules but allowed completion (when not high on marijuana) at personally convenient times. Self-reported data (eg, marijuana use start time, subjective high rating, and quantity) are subject to bias (eg, under- or overreporting). Although mobile cognitive tasks were triggered immediately after the phone survey rating of subjective marijuana high, marijuana effects are short-lived [38], cumulative effects of marijuana use on cognitive functioning could be considered [39], and uncontrolled factors (eg, motivation, other substance use and cause, health conditions such as dyslexia) can impact task performance. The single-item measure of distraction was not well defined and might serve as a proxy for unmeasured contextual influences (eg, presence of others and ambient noise), highlighting the need for improved measurement. Flowers and marijuana Stroop tasks showed possible ceiling effects for number of correct responses. Owing to collinearity and small cell sizes for certain substances, the effects of co-occurring substance use (eg, nicotine) on task performance were not examined. The effects of a person's level of tolerance and cannabis withdrawal on task performance await future research. Correction for multiple comparisons was not done in this pilot study.

Conclusions

Little is known regarding the real-time cognitive impacts resulting from marijuana use in daily life. Although differences

in task performance on the brief mobile cognitive tests when high on marijuana vs not high were small, they were statistically significant and observed for both reaction time and number of correct responses across tasks assessing different cognitive functions. Mobile technology to help detect impacts of acute episodes of marijuana use on cognitive functioning in real time, in the natural environment, could support health care monitoring and provide ongoing feedback to individuals to meet personal health goals [40]. The potential adverse consequences of acute marijuana use on cognitive functioning (eg, while driving) and possible cumulative effects of chronic heavy marijuana use on health compel the development of real-time, mobile methods of monitoring cognitive functioning in the natural environment to help guide personal decision making regarding health behaviors.

Acknowledgments

This study received the following funding: NIDA R21 DA043181, Clinical and Translational Science Award UL1 TTR001857, and National Institute on Alcohol Abuse and Alcoholism R01 AA019511. The funders had no input in the content of the manuscript.

Authors' Contributions

TC contributed to the study design, data collection, data analysis, draft manuscript, and review and approval of the manuscript. SWB contributed to the study design, data collection, data processing, draft manuscript, and review and approval of the manuscript. EYM contributed to the data analysis, draft manuscript, and review and approval of the manuscript. BS contributed to the study design and review and approval of the manuscript. YN contributed to the data collection and review and approval of the manuscript. SJ contributed to the data collection, data processing, and review and approval of the manuscript. AKD contributed to the study design and review and approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Log Likelihood difference testing.
[DOCX File , 983 KB - mhealth_v8i3e16240_app1.docx]

Multimedia Appendix 2
Number of days to complete up to 60 sessions per participant.
[PNG File , 13 KB - mhealth_v8i3e16240_app2.png]

Multimedia Appendix 3
Number of completed sessions per participant.
[PNG File , 13 KB - mhealth_v8i3e16240_app3.png]

References

1. Conroy DA, Kurth ME, Brower KJ, Strong DR, Stein MD. Impact of marijuana use on self-rated cognition in young adult men and women. Am J Addict 2015 Mar;24(2):160-165 [FREE Full text] [doi: 10.1111/ajad.12157] [Medline: 25864605]
2. Volkow ND, Swanson JM, Evins AE, DeLisi LE, Meier MH, Gonzalez R, et al. Effects of cannabis use on human behavior, including cognition, motivation, and psychosis: a review. JAMA Psychiatry 2016 Mar;73(3):292-297. [doi: 10.1001/jamapsychiatry.2015.3278] [Medline: 26842658]
3. Phillips KT, Phillips MM, Lalonde TL, Tormohlen KN. Marijuana use, craving, and academic motivation and performance among college students: an in-the-moment study. Addict Behav 2015 Aug;47:42-47. [doi: 10.1016/j.addbeh.2015.03.020] [Medline: 25864134]
4. Huestis MA. Cannabis-impaired driving: a public health and safety concern. Clin Chem 2015 Oct;61(10):1223-1225. [doi: 10.1373/clinchem.2015.245001] [Medline: 26283691]
5. The National Academies of Science, Engineering, and Medicine, Committee on the Health Effects of Marijuana: An Evidence Review and Research, Board on Population Health and Public Health Practice, Health and Medicine Division. The Health Effects of Cannabis and Cannabinoids: The Current State of Evidence and Recommendations for Research. Washington, DC: National Academies Press; 2017.
6. Scott JC, Slomiak ST, Jones JD, Rosen AF, Moore TM, Gur RC. Association of cannabis with cognitive functioning in adolescents and young adults: a systematic review and meta-analysis. JAMA Psychiatry 2018 Jun 1;75(6):585-595 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.0335] [Medline: 29710074]
7. Zuurman L, Ippel AE, Moin E, van Gerven JM. Biomarkers for the effects of cannabis and THC in healthy volunteers. Br J Clin Pharmacol 2009 Jan;67(1):5-21 [FREE Full text] [doi: 10.1111/j.1365-2125.2008.03329.x] [Medline: 19133057]

8. Crane NA, Schuster RM, Fusar-Poli P, Gonzalez R. Effects of cannabis on neurocognitive functioning: recent advances, neurodevelopmental influences, and sex differences. Neuropsychol Rev 2013 Jun;23(2):117-137 [FREE Full text] [doi: 10.1007/s11065-012-9222-1] [Medline: 23129391]
9. Schreiner AM, Dunn ME. Residual effects of cannabis use on neurocognitive performance after prolonged abstinence: a meta-analysis. Exp Clin Psychopharmacol 2012 Oct;20(5):420-429. [doi: 10.1037/a0029117] [Medline: 22731735]
10. Crean RD, Tapert SF, Minassian A, Macdonald K, Crane NA, Mason BJ. Effects of chronic, heavy cannabis use on executive functions. J Addict Med 2011 Mar;5(1):9-15 [FREE Full text] [doi: 10.1097/ADM.0b013e31820cdd57] [Medline: 21643485]
11. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. Sci Data 2016 Mar 3;3:160011 [FREE Full text] [doi: 10.1038/sdata.2016.11] [Medline: 26938265]
12. Field M. Cannabis 'dependence' and attentional bias for cannabis-related words. Behav Pharmacol 2005 Sep;16(5-6):473-476. [doi: 10.1097/00008877-200509000-00021] [Medline: 16148453]
13. Bettcher BM, Libon DJ, Kaplan E, Swenson R, Penney DL. Digit symbol substitution test. In: Kreutzer JS, DeLuca J, Caplan B, editors. Encyclopedia of Clinical Neuropsychology. New York, NY: Springer; 2011:849-853.
14. Cameron E, Sinclair W, Tiplady B. Validity and sensitivity of a pen computer battery of performance tests. J Psychopharmacol 2001 Jun;15(2):105-110. [doi: 10.1177/026988110101500207] [Medline: 11448083]
15. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol Rev 2003 Dec;13(4):181-197. [doi: 10.1023/b:nerv.0000009483.91468.fb] [Medline: 15000225]
16. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and validity of ambulatory cognitive assessments. Assessment 2018 Jan;25(1):14-30 [FREE Full text] [doi: 10.1177/1073191116643164] [Medline: 27084835]
17. Moore RC, Swendsen J, Depp CA. Applications for self-administered mobile cognitive assessments in clinical research: a systematic review. Int J Methods Psychiatr Res 2017 Dec;26(4) [FREE Full text] [doi: 10.1002/mpr.1562] [Medline: 28370881]
18. Schuster RM, Mermelstein RJ, Hedeker D. Ecological momentary assessment of working memory under conditions of simultaneous marijuana and tobacco use. Addiction 2016 Aug;111(8):1466-1476 [FREE Full text] [doi: 10.1111/add.13342] [Medline: 26857917]
19. Zhang MW, Ying J, Wing T, Song G, Fung DS, Smith HE. Cognitive biases in cannabis, opioid, and stimulant disorders: a systematic review. Front Psychiatry 2018;9:376 [FREE Full text] [doi: 10.3389/fpsyt.2018.00376] [Medline: 30158880]
20. Field M, Munafò MR, Franken IHA. A meta-analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. Psychol Bull 2009 Jul;135(4):589-607 [FREE Full text] [doi: 10.1037/a0015843] [Medline: 19586163]
21. Spanakis P, Jones A, Field M, Christiansen P. A Stroop in the hand is worth two on the laptop: superior reliability of a smartphone based alcohol Stroop in the real world. Subst Use Misuse 2019;54(4):692-698. [doi: 10.1080/10826084.2018.1536716] [Medline: 30572780]
22. Waters AJ, Li Y. Evaluating the utility of administering a reaction time task in an ecological momentary assessment study. Psychopharmacology (Berl) 2008 Mar;197(1):25-35. [doi: 10.1007/s00213-007-1006-6] [Medline: 18064440]
23. Waters AJ, Szeto EH, Wetter DW, Cinciripini PM, Robinson JD, Li Y. Cognition and craving during smoking cessation: an ecological momentary assessment study. Nicotine Tob Res 2014 May;16(Suppl 2):S111-S118 [FREE Full text] [doi: 10.1093/ntr/ntt108] [Medline: 23901053]
24. Bouvard A, Dupuy M, Schweitzer P, Revranche M, Fatseas M, Serre F, et al. Feasibility and validity of mobile cognitive testing in patients with substance use disorders and healthy controls. Am J Addict 2018 Oct;27(7):553-556. [doi: 10.1111/ajad.12804] [Medline: 30260085]
25. 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]
26. Prince J, Arora S, de Vos M. Big data in Parkinson's disease: using smartphones to remotely detect longitudinal disease phenotypes. Physiol Meas 2018 Apr 26;39(4):044005. [doi: 10.1088/1361-6579/aab512] [Medline: 29516871]
27. AWARE Framework. Oulu, Finland URL: https://awareframework.com/ [accessed 2020-02-13]
28. NuRelm. Pittsburgh, PA URL: https://nurelm.com/ [accessed 2020-02-13]
29. Blair JR, Spreen O. Predicting premorbid IQ: a revision of the national adult reading test. Clin Neuropsychol 1989;3(2):129-136. [doi: 10.1080/13854048908403285]
30. NIDA. National Institute on Drug Abuse (NIDA). NIDA Modified ASSIST Bethesda, MD: National Institute on Drug Abuse NIDA Modified ASSIST URL: http://www.nida.nih.gov/nidamed/screening/nmassist.pdf [accessed 2017-12-01]
31. Bright P, Hale E, Gooch VJ, Myhill T, van der Linde I. The National Adult Reading Test: restandardisation against the Wechsler Adult Intelligence Scale-Fourth edition. Neuropsychol Rehabil 2018 Sep;28(6):1019-1027. [doi: 10.1080/09602011.2016.1231121] [Medline: 27624393]
32. Kay M, Rector K, Consolvo S, Greenstein B, Wobbrock J, Watson N, et al. PVT-Touch: Adapting a Reaction Time Test for Touchscreen Devices. In: Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops. New York, NY: IEEE; 2013 Presented at: PERVASIVEHEALTH'13; May 5-8, 2013;

Venice, Italy URL: https://eudl.eu/doi/10.4108/icst.pervasivehealth.2013.252078 [doi: 10.4108/icst.pervasivehealth.2013.252078]
33. Waters AJ, Marhe R, Franken IH. Attentional bias to drug cues is elevated before and during temptations to use heroin and cocaine. Psychopharmacology (Berl) 2012 Feb;219(3):909-921 [FREE Full text] [doi: 10.1007/s00213-011-2424-z] [Medline: 21833505]
34. Hallgren KA, McCrady BS. Interference in the alcohol Stroop task with college student binge drinkers. J Behav Health 2013;2(2):112-119 [FREE Full text] [doi: 10.5455/jbh.20130224082728] [Medline: 25328802]
35. StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC; 2017. URL: https://www.stata.com/support/faqs/resources/citing-software-documentation-faqs/ [accessed 2020-02-05]
36. Hoffman L. Longitudinal Analysis: Modeling Within-Person Fluctuation and Change. New York: Routledge, Taylor & Francis Group; 2015.
37. Prince MA, Conner BT, Pearson MR. Quantifying cannabis: a field study of marijuana quantity estimation. Psychol Addict Behav 2018 Jun;32(4):426-433 [FREE Full text] [doi: 10.1037/adb0000370] [Medline: 29771542]
38. Riedel G, Davies SN. Cannabinoid function in learning, memory and plasticity. Handb Exp Pharmacol 2005(168):445-477. [doi: 10.1007/3-540-26573-2_15] [Medline: 16596784]
39. Pope HG, Gruber AJ, Yurgelun-Todd D. Residual neuropsychologic effects of cannabis. Curr Psychiatry Rep 2001 Dec;3(6):507-512. [doi: 10.1007/s11920-001-0045-7] [Medline: 11707165]
40. Abdullah S, Murnane EL, Matthews M, Kay M, Kientz JA, Gay G, et al. Cognitive Rhythms: Unobtrusive and Continuous Sensing of Alertness Using a Mobile Phone. In: Proceedings of the 2016 ACM International Joint Conference. 2016 Presented at: UbiComp'16; September 12–16, 2016; Heidelberg, Germany URL: https://pac.cs.cornell.edu/pubs/Ubicomp2016_CognitiveRhythms.pdf [doi: 10.1145/2971648.2971712]

Abbreviations

CC: constant-centered
DSST: digit symbol substitution task
EMA: ecological momentary assessment
FSIQ: full-scale IQ
ICC: intraclass correlation
NIDA: National Institute on Drug Abuse
OR: odds ratio

Edited by G Eysenbach; submitted 12.09.19; peer-reviewed by I van der Linde, R Mermelstein; comments to author 07.10.19; revised version received 01.12.19; accepted 28.01.20; published 10.03.20.

Please cite as:
Chung T, Bae SW, Mun EY, Suffoletto B, Nishiyama Y, Jang S, Dey AK
Mobile Assessment of Acute Effects of Marijuana on Cognitive Functioning in Young Adults: Observational Study
JMIR Mhealth Uhealth 2020;8(3):e16240
URL: http://mhealth.jmir.org/2020/3/e16240/
doi:10.2196/16240
PMID:32154789

Original Paper

The Utility of SMS to Report Male Partner HIV Self-testing Outcomes Among Women Seeking Reproductive Health Services in Kenya: Cohort Study

Alison L Drake[1], MPH, PhD; Emily Begnel[1], MPH; Jillian Pintye[2], MPH, PhD; John Kinuthia[3], MBChB, MMed, MPH; Anjuli D Wagner[1], MPH, PhD; Claire W Rothschild[4], MSc; Felix Otieno[3], BA; Valarie Kemunto[3], BScN; Jared M Baeten[1,4,5], MD, MPH, PhD; Grace John-Stewart[1,4,5,6], MD, MPH, PhD

[1]Department of Global Health, University of Washington, Seattle, WA, United States
[2]School of Nursing, University of Washington, Seattle, WA, United States
[3]Research and Programs, Kenyatta National Hospital, Nairobi, Kenya
[4]Department of Epidemiology, University of Washington, Seattle, WA, United States
[5]Department of Medicine, University of Washington, Seattle, WA, United States
[6]Department of Pediatrics, University of Washington, Seattle, WA, United States

Corresponding Author:
Alison L Drake, MPH, PhD
Department of Global Health
University of Washington
325 9th Ave
Seattle, WA
United States
Phone: 1 206 543 5847
Email: adrake2@uw.edu

Abstract

Background: Use of SMS for data collection is expanding, but coverage, bias, and logistical constraints are poorly described.

Objective: The aim of this study is to assess the use of SMS to capture clinical outcomes that occur at home and identify potential biases in reporting compared to in-person ascertainment.

Methods: In the PrEP Implementation in Young Women and Adolescents program, which integrated pre-exposure prophylaxis (PrEP) into antenatal care, postnatal care, and family planning facilities in Kisumu County, Kenya, HIV-negative women 14 years of age or older were offered oral HIV self-tests (HIVSTs) to take home to male partners. Women that brought a phone with a Safaricom SIM to the clinic were offered registration in an automated SMS system (mSurvey) to collect information on HIVST outcomes. Women were asked if they offered the test to their male partners, and asked about the test process and results. HIVST outcomes were collected via SMS (sent 2.5 weeks later), in-person (if women returned for a follow-up scheduled 1 month later), or using both methods (if women initiated PrEP, they also had scheduled follow-up visits). The SMS prompted women to reply at no charge. HIVST outcomes were compared between women with scheduled follow-up visits and those without (follow-up visits were only scheduled for women who initiated PrEP). HIVST outcomes were also compared between women reporting via SMS and in-person.

Results: Among 2123 women offered HIVSTs and mSurvey registration, 486 (23.89%) accepted HIVSTs, of whom 359 (73.87%) were eligible for mSurvey. Additionally, 76/170 (44.7%) women with scheduled follow-up visits and 146/189 (77.3%) without scheduled follow-up visits registered in mSurvey. Among the 76 women with scheduled follow-ups, 62 (82%) had HIVST outcomes collected: 19 (31%) in-person, 20 (32%) by SMS, and 23 (37%) using both methods. Among the 146 women without scheduled visits, 87 (59.6%) had HIVST outcomes collected: 3 (3%) in-person, 82 (94%) by SMS, and 2 (2%) using both methods. SMS increased the collection of HIVST outcomes substantially for women with scheduled follow-up visits (1.48-fold), and captured 82 additional reports from women without scheduled follow-up visits. Among 222 women with reported HIVST outcomes, frequencies of offering partners the HIVST (85/95, 89% in-person vs 96/102, 94% SMS; P=.31), partners using the HIVST (83/85, 98% vs 92/96, 96%; P=.50), women using HIVST with partners (82/83, 99% vs 91/92, 99%; P=.94), and seeing partner's HIVST results (82/83, 99% vs 89/92, 97%; P=.56) were similar between women reporting in-person only versus by SMS only.

However, frequency of reports of experiencing harm or negative reactions from partners was more commonly reported in the SMS group (17/102, 16.7% vs 2/85, 2%; *P*=.003). Barriers to the SMS system registration included not having a Safaricom SIM or a functioning phone.

Conclusions: Our results suggest that the use of SMS substantially improves completeness of outcome data, does not bias reporting of nonsensitive information, and may increase reporting of sensitive information.

(JMIR Mhealth Uhealth 2020;8(3):e15281) doi:10.2196/15281

KEYWORDS

SMS; HIV self-testing; survey coverage; HIV pre-exposure prophylaxis

Introduction

SMS has enormous potential for public health. This technology is inexpensive and has increasing reach with expanding global cellular network coverage and phone ownership. SMS has been used to provide reminders [1,2], return lab results [3], provide education and improve knowledge [4], and promote healthy behaviors [5,6]. SMS has also been used to remotely collect survey data on health outcomes and behaviors, a strategy that may reduce the travel time to and cost of follow-up visits, and allows participation at convenient times [7].

Studies in low- and middle-income countries comparing SMS to other data collection approaches or offering choices in approaches are limited [8-10]. Some individuals may be able to overcome barriers to in-person visits, or may find that using both SMS and in-person approaches to survey assessment is acceptable and feasible. However, SMS may also be useful for individuals who might otherwise decline participation or become lost to follow-up. Self-administered surveys may reduce social desirability bias [11-13], although differential outcome ascertainment may bias results by using multiple approaches to data collection. Using a combination of strategies to capture health outcomes may improve participation and generalizability, but it is important to measure outcomes using different strategies in the same context and setting to determine whether results are biased based on the strategy used. We measured the utility of incorporating SMS as an alternative, complementary strategy to in-person assessment of male partner HIV self-test (HIVST) outcomes by women, and assessed bias in reporting results using either method.

Methods

From November 20, 2017, to June 15, 2018, 3425 women seeking antenatal care (ANC), postpartum care (PNC), or family planning (FP) services at 8 facilities in Kisumu County, Kenya were asked to take an HIVST home to their male partners as part of a pre-exposure prophylaxis (PrEP) implementation program. Women were offered PrEP, and those that accepted had a PrEP follow-up visit scheduled 1 month later where HIVST outcomes were ascertained in-person. Women were classified as having scheduled follow-up visits if they initiated PrEP. Women who declined PrEP but still took an HIVST had outcomes ascertained in-person if they returned for maternal and child health or family planning services [14].

A subset of women who agreed to take an HIVST were asked to register in an automated SMS communication system (mSurvey; Nairobi, Kenya) to assess HIVST outcomes. Registration was offered in all program sites over time (initially only offered by 1 site, but expanded to all 8 sites by the end of program activity) for eligible women. Women were eligible to register if they had their phone at the clinic and a Safaricom SIM card, and provided oral consent. Registered women selected their preferred language (English, Kiswahili, or Dholou) and were asked to save the mSurvey phone number as a contact in their phone to ensure the phone number would be recognized when the follow-up survey was sent (2.5 weeks later). Women were informed that responding to the mSurvey SMS was free, and they could opt-out at any time. HIVST outcomes were assessed through sequential SMS inquiries sent by mSurvey for women to respond to using numerical responses representing survey answers. Women who initiated registration or follow-up surveys were given 72 hours to complete surveys before they were timed-out of the system and unable to complete the survey.

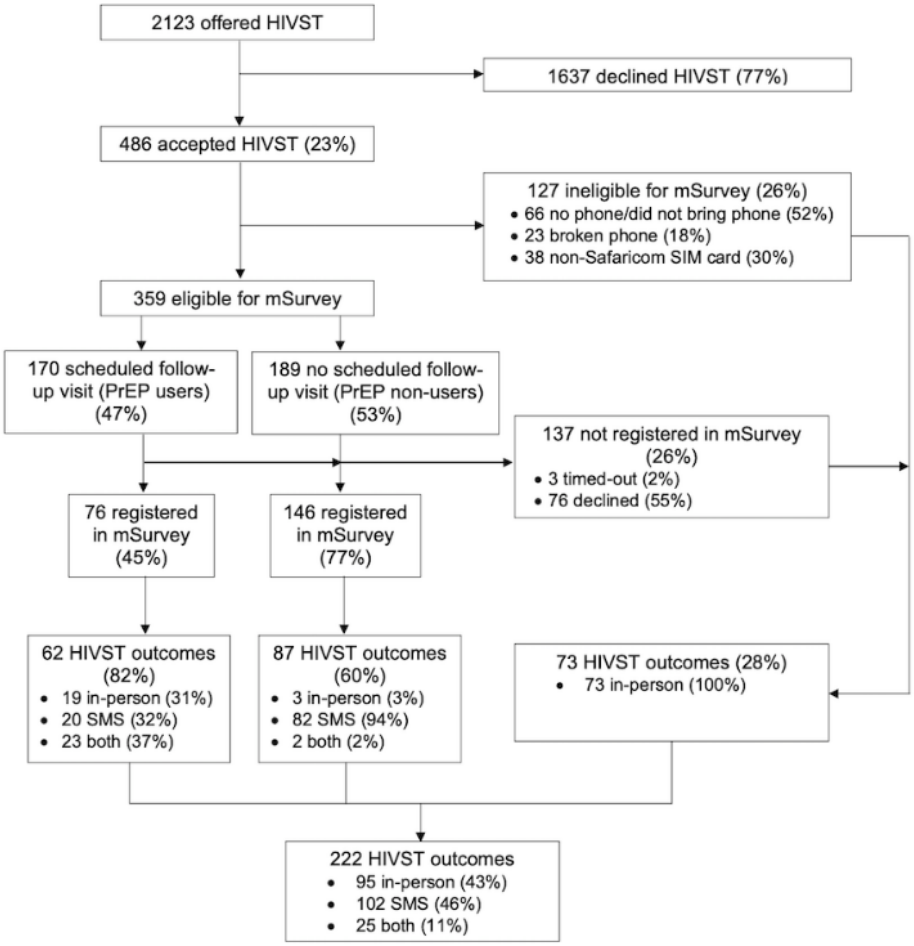
Mode of HIVST outcome ascertainment (SMS vs in-person) was compared between women with and without scheduled follow-up visits. HIVST outcomes were compared between SMS and in-person responders. The primary analysis was restricted to women offered HIVSTs at facilities when mSurvey registration was available. Women who reported having a partner with HIV at enrollment were excluded from the analysis. Continuous and categorical variables were compared using the Wilcoxon rank-sum and chi-square tests, respectively. Statistical analyses were performed using Stata v14 (StataCorp LLC; College Station, TX).

This study was reviewed and approved by the University of Washington Institutional Review Board and the Kenyatta National Hospital Ethics & Research Committee.

Results

In total, 486 women accepted HIVSTs; 359 were eligible for registration in the mSurvey SMS system, and 222 successfully registered (Figure 1). Some women were ineligible due to the logistical barriers of not having a phone, not having a working phone, or not having a Safaricom SIM card. Nearly half (98/222, 43.6%) of the women who successfully registered for mSurvey timed-out before completing the follow-up SMS survey.

Figure 1. Flow diagram of women accepting HIV self-tests for male partners. HIVST: HIV self-test; PrEP: pre-exposure prophylaxis.



Overall, among 222 women enrolled in mSurvey: 116 (52.3%) were from ANC, 80 (36.0%) were from PNC, and 26 (11.7%) were from FP. The median age was 25 (IQR 22-28); 86.4% (191/222) were married and 17.1% (38/222) knew their male partners were HIV-negative. Intimate partner violence (IPV) was experienced by 4.5% (10/222) of the enrolled women within the last 6 months. HIVST outcomes were more likely to be reported through SMS only (vs in-person only), by women who were older (median 25 vs 23 years of age; *P*=.01), or if they had a partner of unknown HIV status (92/102, 90.2% vs 72/95, 76%; *P*=.01). Marital status (*P*=.84), history of IPV (*P*=.16), and risk factors for HIV (transactional sex, *P*=.52; diagnosis with or treatment for STI, *P*=0.14; forced sex, *P*=0.14; shared needles while engaging in intravenous drug use, *P*=0.14; and used post-exposure prophylaxis more than twice, *P*=0.14) were similar between women with HIVST outcomes assessed in-person and via SMS.

Of the 76 women registered in mSurvey with scheduled follow-up visits, 82% (62) had HIVST outcome data available. SMS increased outcome ascertainment 1.48-fold (relative risk; CI 1.32-1.64); an additional 32% of HIVST outcomes would have been missed without SMS. The majority of women enrolled in mSurvey but without scheduled follow-up visits (82/87, 94%) reported HIVST outcomes by SMS only. SMS responses captured 102/149 (68.5%) of all HIVST outcomes assessed by women in mSurvey.

HIVST outcomes were also assessed for 73 women who were ineligible or not registered in mSurvey. The frequency of reporting on HIVST outcomes by women with scheduled follow-up visits who were ineligible (32/46, 70%) or who did not register in mSurvey (41/94, 44%) was similar to the frequency of women who registered in mSurvey. Most HIVST outcomes were reported near the time SMS surveys were sent to women using SMS (median 0 days after SMS delivered, maximum 46 days), or near the scheduled follow-up date for women reporting in-person (median 0 days after scheduled date, IQR –2 to 5, range –29 to 141).

Table 1 shows the data collected from reports on HIVST experiences. The offer of HIVSTs to partners was similar between women reporting in-person and via SMS. There were no differences in reporting between in-person and SMS regarding whether or not partners took the test and if the women saw the results. A history of IPV was lower in women who reported HIVST outcomes by SMS vs in-person (3/102, 2.9% vs 7/95, 7%; *P*=.13). However, reports of experiencing harm or negative reactions from partners as a result of the HIVST were significantly more frequent in women with HIVST outcomes assessed via SMS than in-person.

Table 1. Partner HIV self-testing experience reports.

	Overall, n (%)	In-person, n (%)	SMS, n (%)	Both, n (%)	*P* value[a]
Offered partner HIVST [b]	N=222	N=95	N=102	N=25	.31
No	15 (6.8)	9 (9.5)	5 (4.9)	1 (4.0)	
Yes	205 (92.3)	85 (89.5)	96 (94.1)	24 (96.0)	
Refused to answer/Don't know	2 (0.5)	1 (1.1)	1 (1.0)	0 (0.0)	
Reasons partner HIVST not offered	N=10	N=9	—[c]	N=1	—
Fear of partner's reaction	1 (10.0)	0 (0.0)	—	1 (100.0)	
Have not seen partner since HIVST received	3 (30.0)	3 (33.3)	—	0 (0.0)	
Tried to discuss, partner reacted negatively	4 (40.0)	4 (44.4)	—	0 (0.0)	
Other	2 (20.0)	2 (22.2)	—	0 (0.0)	
Partner used HIVST	N=205	N=85	N=96	N=24	.50
	199 (97.1)	83 (97.7)	92 (95.8)	1 (100.0)	
Tested with partner	N=199	N=83	N=92	N=24	.94
	197 (99.0)	82 (98.8)	91 (98.9)	24 (100.0)	
Saw partner's results	N=199	N=83	N=92	N=24	.56
Yes, I observed it	195 (98.0)	82 (98.8)	89 (96.7)	24 (100.0)	
No, I was told the results	3 (1.5)	1 (1.2)	2 (2.2)	0 (0.0)	
No, I don't know his result	1 (0.5)	0 (0.0)	1 (1.1)	0 (0.0)	
Partner HIV results	N=198	N=83	N=91	N=24	.52
HIV-negative	190 (96.0)	80 (96.4)	87 (95.6)	23 (95.8)	
HIV-positive	4 (2.0)	2 (2.4)	1 (1.1)	1 (4.2)	
Refused to answer	2 (1.0)	1 (1.2)	1 (1.1)	0 (0.0)	
Missing	2 (1.0)	0 (0.0)	2 (2.2)	0 (0.0)	
Experienced harm as a result of the HIVST	N=222	N=95	N=102	N=25	<.01
No	190 (85.6)	85 (89.5)	80 (78.4)	25 (100.0)	
Yes	19 (8.6)	2 (2.1)	17 (16.7)	0 (0.0)	
Refused to answer	1 (0.4)	0 (0.0)	1 (1.0)	0 (0.0)	
Missing	12 (5.4)	8 (8.4)	4 (3.9)	0 (0.0)	

[a]Comparing in-person and SMS groups.

[b]HIVST: HIV self-test.

[c]SMS reporting results not evaluated, and *P* value not determined.

Discussion

We found that nearly half of women seeking reproductive health services were willing and able to use SMS to respond to surveys on sensitive HIVST outcomes for their male partners. The SMS-based survey substantially increased the proportion of women that reported HIVST outcomes among women with scheduled follow-up visits (1.48-fold increase), and was a successful strategy to capture HIVST outcomes for those without scheduled follow-up visits. We did not detect differences in the proportions of women who reported that partners used HIVSTs or reported their partner's HIV status; however, we did note a significant difference in the frequency of reported social harm related to HIVSTs. This suggests that the mode of outcome ascertainment did not bias responses for nonsensitive questions and may have improved the reporting of sensitive information, such as reports of social harm related to HIVSTs.

Our analysis was subject to some limitations, but also provided insight on logistical barriers to SMS registration. Restricting registration to the primary mobile carrier and requiring phones to be present led to exclusions of 21.3% (104/486) of women accepting HIVSTs. We were unable to disaggregate lack of phone ownership from phones not being brought to the clinic. Women in our study may also have had intermittent access to a mobile phone due to shared phones with male partners or other community members, which have previously been reported as potential barriers to including women in SMS surveys [15]. System time-out at registration was uncommon, but 43.6% (96/222) timed-out at follow-up and did not complete the survey. These findings may indicate poorer network coverage and power

supply problems outside the clinic, or that participants became disinterested or uncomfortable answering the SMS surveys. Mobile phone accessibility, network coverage consistency, and difference in mobile carriers may impact generalizability and should be considered during the development stage of SMS projects [15]. Alternative strategies for survey registration, such as remote registration for individuals who share phones or do not have their phone present, may improve eligibility for SMS surveys. Finally, this study may have lacked power to compare some HIVST outcomes between women responding via SMS and in-person, such as partners with HIV, and was not able to compare reasons for not offering HIVSTs since this information was not collected via SMS.

SMS has successfully been used to remotely collect survey data [16-20]. Studies suggest that participants are willing to respond to sensitive questions if reminders to delete messages, passwords, or personal identification numbers are used [21,22]. In East Africa, SMS response rates to surveys on sexual behaviors, pregnancy history, HIV testing, and adherence to PrEP range between 14% and 96% [16,21,23-25]. Response rates increase with financial incentives, clear instructions for responding, and in the context of research [17,21,26]. Lack of incentives coupled with misconceptions that costs would be incurred for SMS responses may have hindered response rates in our program compared to other studies [27,28].

In conclusion, SMS enhanced our ability to measure male partner HIVST outcomes. Our study demonstrates that SMS can be used to collect brief survey data on sensitive information, even in the absence of financial incentives. SMS should be considered to capture health outcomes, which may alleviate health system constraints and burdens associated with in-person visits.

Acknowledgments

The PrEP Implementation for Young Women and Adolescents Program was funded by the United States Department of State as part of the DREAMS Innovation Challenge (Grant #37188-1088 MOD01), managed by JSI Research & Training Institute, Inc. Support for the project was also received by NIH/NIAID K01 AI116298 (ALD), NIH/NICHD F32HD088204 (ADW), the University of Washington (UW)/Fred Hutch Center for AIDS Research (CFAR) NIH/NIAID P30-AI027757, and the UW Global Center for Integrated Health of Women, Adolescents, and Children (Global WACh).

Conflicts of Interest

None declared.

References

1. Boksmati N, Butler-Henderson K, Anderson K, Sahama T. The effectiveness of SMS reminders on appointment attendance: a meta-analysis. J Med Syst 2016 Apr;40(4):90. [doi: 10.1007/s10916-016-0452-2] [Medline: 26852337]
2. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]
3. Seidenberg P, Nicholson S, Schaefer M, Semrau K, Bweupe M, Masese N, et al. Early infant diagnosis of HIV infection in Zambia through mobile phone texting of blood test results. Bull World Health Organ 2012 May 01;90(5):348-356 [FREE Full text] [doi: 10.2471/BLT.11.100032] [Medline: 22589568]
4. Johnson D, Juras R, Riley P, Chatterji M, Sloane P, Choi SK, et al. A randomized controlled trial of the impact of a family planning mHealth service on knowledge and use of contraception. Contraception 2017 Jan;95(1):90-97. [doi: 10.1016/j.contraception.2016.07.009] [Medline: 27421767]
5. Odeny TA, Bailey RC, Bukusi EA, Simoni JM, Tapia KA, Yuhas K, et al. Effect of text messaging to deter early resumption of sexual activity after male circumcision for HIV prevention: a randomized controlled trial. J Acquir Immune Defic Syndr 2014 Feb 01;65(2):e50-e57 [FREE Full text] [doi: 10.1097/QAI.0b013e3182a0a050] [Medline: 23846561]
6. Unger JA, Ronen K, Perrier T, DeRenzi B, Slyker J, Drake AL, et al. Short message service communication improves exclusive breastfeeding and early postpartum contraception in a low- to middle-income country setting: a randomised trial. BJOG 2018 Nov;125(12):1620-1629 [FREE Full text] [doi: 10.1111/1471-0528.15337] [Medline: 29924912]
7. Goel S, Bhatnagar N, Sharma D, Singh A. Bridging the human resource gap in primary health care delivery systems of developing countries with mHealth: narrative literature review. JMIR Mhealth Uhealth 2013 Dec 03;1(2):e25 [FREE Full text] [doi: 10.2196/mhealth.2688] [Medline: 25099436]
8. Musinguzi N, Muwonge T, Ngure K, Katabira E, Mugo N, Burns BFO, Partners Mobile Adherence to PrEP (PMAP) Team. Comparison of short messaging service self-reported adherence with other adherence measures in a demonstration project of HIV preexposure prophylaxis in Kenya and Uganda. AIDS 2018 Sep 24;32(15):2237-2245 [FREE Full text] [doi: 10.1097/QAD.0000000000001955] [Medline: 30005025]
9. Ballivian A, Azevedo J, Durbin W, Rios J, Godoy J, Borisova C, et al. Listening to LAC: Using Mobile Phones for High Frequency Data Collection: Final Report Washington DC. Washington DC: The World Bank; 2013 Feb. URL: http://catalog.ihsn.org/index.php/catalog/4775/related_materials [accessed 2020-01-27]

10. Gibson DG, Pereira A, Farrenkopf BA, Labrique AB, Pariyo GW, Hyder AA. Mobile phone surveys for collecting population-level estimates in low- and middle-income countries: a literature review. J Med Internet Res 2017 May 05;19(5):e139 [FREE Full text] [doi: 10.2196/jmir.7428] [Medline: 28476725]
11. Bruun S, Buhl S, Husby S, Jacobsen LN, Michaelsen KF, Sørensen J, et al. Breastfeeding, infant formula, and introduction to complementary foods-comparing data obtained by questionnaires and health visitors' reports to weekly short message service text messages. Breastfeed Med 2017 Nov;12(9):554-560. [doi: 10.1089/bfm.2017.0054] [Medline: 28832183]
12. Kelly CA, Soler-Hampejsek E, Mensch BS, Hewett PC. Social desirability bias in sexual behavior reporting: evidence from an interview mode experiment in rural Malawi. Int Perspect Sex Reprod Health 2013 Mar;39(1):14-21 [FREE Full text] [doi: 10.1363/3901413] [Medline: 23584464]
13. Langhaug LF, Sherr L, Cowan FM. How to improve the validity of sexual behaviour reporting: systematic review of questionnaire delivery modes in developing countries. Trop Med Int Health 2010 Mar;15(3):362-381 [FREE Full text] [doi: 10.1111/j.1365-3156.2009.02464.x] [Medline: 20409291]
14. Pintye J, Drake AL, Begnel E, Kinuthia J, Abuna F, Lagat H, et al. Acceptability and outcomes of distributing HIV self-tests for male partner testing in Kenyan maternal and child health and family planning clinics. AIDS 2019 Jul 01;33(8):1369-1378 [FREE Full text] [doi: 10.1097/QAD.0000000000002211] [Medline: 30932954]
15. Kumoji E, Khan Sohail S. Use of short message service for monitoring Zika-related behaviors in four Latin American countries: lessons learned from the field. Mhealth 2019;5:23 [FREE Full text] [doi: 10.21037/mhealth.2019.07.01] [Medline: 31559268]
16. Muwonge TR, Ngure K, Katabira E, Mugo N, Kimemia G, Burns BFO, Partners Mobile Adherence to PrEP (PMAP) Team. Short Message Service (SMS) Surveys Assessing Pre-exposure Prophylaxis (PrEP) Adherence and Sexual Behavior are Highly Acceptable Among HIV-Uninfected Members of Serodiscordant Couples in East Africa: A Mixed Methods Study. AIDS Behav 2019 May;23(5):1267-1276. [doi: 10.1007/s10461-018-2326-8] [Medline: 30406335]
17. Leidich A, Jayaweera R, Arcara J, Clawson S, Chalker C, Rochat R. Evaluating the feasibility and acceptability of sending pregnancy and abortion history surveys through SMS text messaging to help reach sustainable development goal 3. Int J Med Inform 2018 Jun;114:108-113. [doi: 10.1016/j.ijmedinf.2017.10.017] [Medline: 29100753]
18. Demombynes G, Gubbins P, Romeo A. Policy Research Working Paper. 2013 Jan. Challenges and Opportunities of Mobile Phone-Based Data Collection: Evidence from South Sudan URL: http://documents.worldbank.org/curated/en/385371468102891764/Challenges-and-opportunities-of-mobile-phone-based-data-collection-evidence-from-South-Sudan [accessed 2020-01-27]
19. de Tolly KM, Constant D. Integrating mobile phones into medical abortion provision: intervention development, use, and lessons learned from a randomized controlled trial. JMIR Mhealth Uhealth 2014 Mar 14;2(1):e5 [FREE Full text] [doi: 10.2196/mhealth.3165] [Medline: 25098569]
20. Vahdat HL, L'Engle KL, Plourde KF, Magaria L, Olawo A. There are some questions you may not ask in a clinic: providing contraception information to young people in Kenya using SMS. Int J Gynaecol Obstet 2013 Nov;123 Suppl 1:e2-e6 [FREE Full text] [doi: 10.1016/j.ijgo.2013.07.009] [Medline: 24012514]
21. Haberer JE, Kiwanuka J, Nansera D, Wilson IB, Bangsberg DR. Challenges in using mobile phones for collection of antiretroviral therapy adherence data in a resource-limited setting. AIDS Behav 2010 Dec;14(6):1294-1301 [FREE Full text] [doi: 10.1007/s10461-010-9720-1] [Medline: 20532605]
22. Siedner MJ, Haberer JE, Bwana MB, Ware NC, Bangsberg DR. High acceptability for cell phone text messages to improve communication of laboratory results with HIV-infected patients in rural Uganda: a cross-sectional survey study. BMC Med Inform Decis Mak 2012 Jun 21;12:56 [FREE Full text] [doi: 10.1186/1472-6947-12-56] [Medline: 22720901]
23. Curran K, Mugo NR, Kurth A, Ngure K, Heffron R, Donnell D, et al. Daily short message service surveys to measure sexual behavior and pre-exposure prophylaxis use among Kenyan men and women. AIDS Behav 2013 Nov;17(9):2977-2985 [FREE Full text] [doi: 10.1007/s10461-013-0510-4] [Medline: 23695519]
24. Johnson D. Collecting data from mHealth users via SMS surveys: a case study in Kenya. Surv Pract 2016 Feb 01;9(1):1-8. [doi: 10.29115/sp-2016-0004]
25. Chib A, Wilkin H, Ling LX, Hoefman B, Van Biejma H. You have an important message! Evaluating the effectiveness of a text message HIV/AIDS campaign in Northwest Uganda. J Health Commun 2012;17 Suppl 1:146-157. [doi: 10.1080/10810730.2011.649104] [Medline: 22548607]
26. Gibson DG, Ochieng B, Kagucia EW, Were J, Hayford K, Moulton LH, et al. Mobile phone-delivered reminders and incentives to improve childhood immunisation coverage and timeliness in Kenya (M-SIMU): a cluster randomised controlled trial. Lancet Glob Health 2017 Dec;5(4):e428-e438 [FREE Full text] [doi: 10.1016/S2214-109X(17)30072-4] [Medline: 28288747]
27. Déglise C, Suggs LS, Odermatt P. Short message service (SMS) applications for disease prevention in developing countries. J Med Internet Res 2012 Jan 12;14(1):e3 [FREE Full text] [doi: 10.2196/jmir.1823] [Medline: 22262730]
28. 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]

Abbreviations

ANC: antenatal care
FP: family planning
HIVST: HIV self-test
IPV: intimate partner violence
PNC: postpartum care
PrEP: pre-exposure prophylaxis.

Edited by G Eysenbach; submitted 28.06.19; peer-reviewed by B Smith, L Sun; comments to author 01.10.19; revised version received 16.10.19; accepted 19.12.19; published 25.03.20.

Please cite as:
Drake AL, Begnel E, Pintye J, Kinuthia J, Wagner AD, Rothschild CW, Otieno F, Kemunto V, Baeten JM, John-Stewart G
The Utility of SMS to Report Male Partner HIV Self-testing Outcomes Among Women Seeking Reproductive Health Services in Kenya: Cohort Study
JMIR Mhealth Uhealth 2020;8(3):e15281
URL: http://mhealth.jmir.org/2020/3/e15281/
doi:10.2196/15281
PMID:32209530

Original Paper

Volumetric Food Quantification Using Computer Vision on a Depth-Sensing Smartphone: Preclinical Study

David Herzig[1*], PhD; Christos T Nakas[2,3*], PhD; Janine Stalder[1], MSc; Christophe Kosinski[1], MD; Céline Laesser[1], MD; Joachim Dehais[1], PhD; Raphael Jaeggi[1]; Alexander Benedikt Leichtle[3], MD; Fried-Michael Dahlweid[4], PhD; Christoph Stettler[1*], MD; Lia Bally[1*], MD, PhD

[1]Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism, Bern University Hospital, University of Bern, Bern, Switzerland

[2]Laboratory of Biometry, University of Thessaly, Nea Ionia/Volos, Magnesia, Greece

[3]University Institute of Clinical Chemistry, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

[4]Insel Data Science Center & Department for Technology and Innovation, Bern University Hospital, University of Bern, Bern, Switzerland

[*]these authors contributed equally

Corresponding Author:
Lia Bally, MD, PhD
Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism
Bern University Hospital
University of Bern
Freiburgstrasse 15
Bern, 3010
Switzerland
Phone: 41 31 632 36 77
Email: lia.bally@insel.ch

Abstract

Background: Quantification of dietary intake is key to the prevention and management of numerous metabolic disorders. Conventional approaches are challenging, laborious, and lack accuracy. The recent advent of depth-sensing smartphones in conjunction with computer vision could facilitate reliable quantification of food intake.

Objective: The objective of this study was to evaluate the accuracy of a novel smartphone app combining depth-sensing hardware with computer vision to quantify meal macronutrient content using volumetry.

Methods: The app ran on a smartphone with a built-in depth sensor applying structured light (iPhone X). The app estimated weight, macronutrient (carbohydrate, protein, fat), and energy content of 48 randomly chosen meals (breakfasts, cooked meals, snacks) encompassing 128 food items. The reference weight was generated by weighing individual food items using a precision scale. The study endpoints were (1) error of estimated meal weight, (2) error of estimated meal macronutrient content and energy content, (3) segmentation performance, and (4) processing time.

Results: In both absolute and relative terms, the mean (SD) absolute errors of the app's estimates were 35.1 g (42.8 g; relative absolute error: 14.0% [12.2%]) for weight; 5.5 g (5.1 g; relative absolute error: 14.8% [10.9%]) for carbohydrate content; 1.3 g (1.7 g; relative absolute error: 12.3% [12.8%]) for fat content; 2.4 g (5.6 g; relative absolute error: 13.0% [13.8%]) for protein content; and 41.2 kcal (42.5 kcal; relative absolute error: 12.7% [10.8%]) for energy content. Although estimation accuracy was not affected by the viewing angle, the type of meal mattered, with slightly worse performance for cooked meals than for breakfasts and snacks. Segmentation adjustment was required for 7 of the 128 items. Mean (SD) processing time across all meals was 22.9 seconds (8.6 seconds).

Conclusions: This study evaluated the accuracy of a novel smartphone app with an integrated depth-sensing camera and found highly accurate volume estimation across a broad range of food items. In addition, the system demonstrated high segmentation performance and low processing time, highlighting its usability.

(JMIR Mhealth Uhealth 2020;8(3):e15294) doi:10.2196/15294

KEYWORDS

depth camera; computer vision; dietary assessment; smartphone

Introduction

Qualitative and quantitative assessment of dietary intake are cornerstones for the prevention and management of metabolic diseases such as obesity and diabetes [1,2]. Traditional manual food records that rely on human abilities to quantify food intake are time-consuming and error-prone [3]. One of the main challenges is the appropriate estimation of portion size (ie, volume) [4]. Inaccurate portion size estimation contributes up to 50% of the total estimation error [5]. Novel approaches replacing manual input by automated techniques may overcome the inherent limitations of traditional approaches, while increasing usability.

Mobile devices, currently ubiquitous, could simplify dietary monitoring. Although there are a number of commercially available apps offering access to food composition databases or providing reference images to facilitate estimation of portion size [6], they are generally limited by the need for manual user input.

High-quality smartphone cameras and computer vision approaches can be combined to fully automate portion size estimation. Users capture images of the meal using the smartphone camera, and the app subsequently builds a 3D model of the food to calculate its volume [7]. Combining the food volume with macronutrient-density databases, the app translates the volume into weight and then nutrient information. Food identification can be accomplished either by user selection or as part of the automated image processing, which further minimizes the need for user input [8].

Researchers have described several such systems [9-13]. A major challenge in many of these approaches lies in the capturing of the third dimension (depth) due to geometric constraints. In particular, factors such as precise food location, shape and size of food items, and changes in these parameters depending on camera perspective potentially interfere with reliable depth assessment. To overcome such constraints, fiducial markers, which ground the scene in a common frame of reference, are utilized. In addition, some systems use multiple images or video sequences of the food, followed by a complex calibration process. All these aspects inherently affect usability and accuracy.

The recent advent of miniaturized depth-sensing cameras embedded within smartphones (eg, iPhone X) opens a new horizon for automated food quantification. Using a single capture including depth information from any convenient viewing angle, this technology has the potential to eliminate the need for manual input, thereby increasing usability as well as accuracy. Therefore, the aim of this study was to evaluate the accuracy of a novel smartphone app that combines depth-sensing with computer vision to quantify food volume across a broad range of meals reflecting a real-life setting.

Methods

Study Design

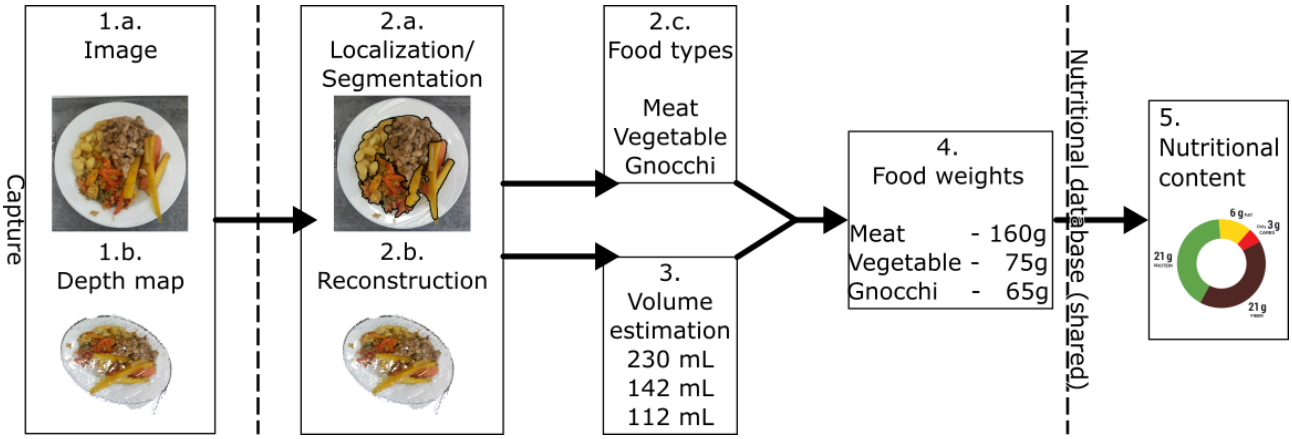
The study occurred at the Central Kitchen Facility of the University Hospital Bern, Switzerland, in mid-January 2019. The system was tested on regular meals served to patients and hospital staff. A total of 48 test meals were randomly generated from a pre-defined pool of 128 food items. The test meals comprised the following meal types: breakfast, cooked meals, and snacks. Meals consisted of 4 (breakfast), 3 (cooked meals), or 1 (snacks) food items and were served on a standard plate or in a standard bowl. The meal content is described in Multimedia Appendix 1.

For each test meal, a single image was taken at a predefined angle of either 45° or 90° from the horizontal position (the angle was estimated by the user). The allocation of the capture angle of each meal was pre-defined using a balanced randomization procedure. The randomization sequence was produced as a binary sequence in three batches (by meal type). Images were captured under natural light conditions.

Smartphone App

The app was provided by SNAQ GmbH (Zurich, Switzerland) using a software version from May 2018, and it was installed on an iPhone X (Apple Inc, Cupertino, California), which uses a built-in depth sensor applying structured light. The automated food quantification workflow is summarized in Figure 1 and consisted of (1) capturing the scene, (2) analyzing the scene, (3) estimating the volume, (4) converting the food volume into food weight, and (5) conversion of the food weight into macronutrient content.

Figure 1. Automated food quantification workflow performed by the SNAQ app.



First, the user takes a photograph using the phone, and a depth map of the food is generated through the phone's front sensors consisting of a photo camera and an active depth sensor.

Second, the system partitions the image into consistent regions representing different items and eliminates those that are not food. To do so, a convolutional neural network has learned how food is structured in terms of sets of pixels and their correlations to the visual appearance of images. The data used to train the system consist of images with flags for each pixel indicating whether the respective pixel represents food. If the automated segmentation is not deemed satisfactory, the user can manually adjust the outlines of the items. A workflow of the segmentation as well as an example of good and bad segmentations are provided in Multimedia Appendix 2. Then, based on the depth map and input from the phone sensors, the visible point cloud is transformed into a set of surfaces using a Delaunay triangulation. The system extracts the location and orientation of the table (vertical plane) using the RANSAC-Algorithm [14] for an outlier-robust fitting. From this, the surfaces of each dish are defined. Selection of the food type is manually performed for each of the segmented food items.

Third, the segmented food items are used to cut the visible surface into partial food surfaces. Each food surface is then closed by the dish surfaces before their volume is calculated.

Fourth, the food volume is converted into food weight using a food density database. Finally, the food weight is converted into macronutrient content using the Swiss Food Composition Database [15].

This study was designed to assess the accuracy of automated quantification of portion sizes. Automated food recognition (ie, the taxonomy of the food items) was not a focus of this study. Instead, the user capturing the image selected the respective food item from a pre-defined list within the app.

Reference Method

The reference weight was generated by weighing individual food items to the nearest 0.1 g using a precision scale (ME4002, Mettler Toledo, Greifensee, Switzerland). Conversion into macronutrient content was performed using the Swiss Food Composition Database [15].

Endpoints

There were four study endpoints: (1) error of the estimated meal weight, (2) error of the estimated meal macronutrient content and energy content, (3) segmentation performance (defined as the number of items requiring manual correction of segmentation as well as intersection of the uncorrected and corrected segmentation areas over the corrected segmentation area), and (4) processing time (defined as the time period from image capture to macronutrient/energy output, including the time required for manual inputs).

Sample Size Calculation and Statistical Analysis

The number of test meals in this study was determined based on a pilot experiment showing a mean (SD) difference in carbohydrate content of –2.6 g (9.2 g). Applying a power of 80% and significance level of .05 for 48 meals was deemed appropriate.

The error was determined on the meal level, and the following error metrics were used: bias, defined as the difference between estimation and reference (estimate-reference); absolute error, defined as |estimate-reference|; and 95% limits of agreement, calculated as ±2*SD of the bias. Bland-Altman plots were generated to visualize the level of agreement between the estimate and reference values. General linear models were used to assess the effect of meal type and inclination angle on the estimation error. *P* values <.05 were considered statistically significant. SPSS version 25.0 (IBM Corp, Armonk, NY) was used for statistical analysis. Data are described using mean (SD) and median (interquartile range [IQR]). All absolute error and bias values in this paper are presented as absolute values (g) followed by the relative values (%) in parentheses.

Results

Macronutrient and Energy Content of the Test Meals

The 48 test meals encompassed 128 food items. The mean reference macronutrient and energy contents of the 48 test meals are summarized in Multimedia Appendix 3. On average, the meals weighed 235.8 g (range 29.6-582.4 g). Meals contained an average 38.5 g carbohydrate (range 4.4-101.0 g), 14.6g protein (range 0.2-66.9 g), and 11.7 g fat (range 0.1-37.1 g). Mean energy content was 325 kcal (range 32-609 kcal). Insights into the study meals, including the representation of different meal types, are provided in Multimedia Appendix 4.

Errors of Estimated Meal Weight, Estimated Meal Macronutrient Content, and Estimated Meal Energy Content

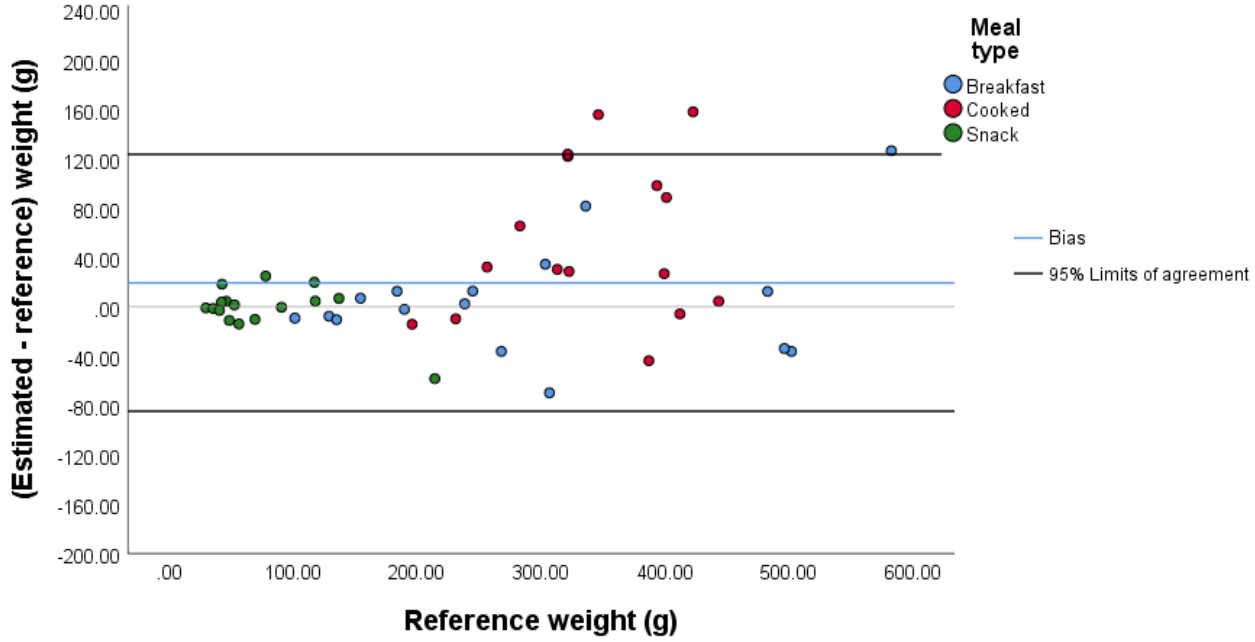
The mean (SD) error metrics are summarized in Table 1, and the median (IQR) error metrics are presented in Multimedia Appendix 5. Corresponding Bland-Altman plots are presented in Figures 2-6. In both absolute and relative terms, the mean (SD) absolute error of the estimated weight for all meals was 35.1 g (42.8 g; 14.0% [12.2%]), and the mean (SD) bias was 19.3 g (52.1 g; 5.4% [17.8%]). The 95% limits of agreement were –84.8 g and 123.4 g (Figure 2).

Table 1. Error metrics for all meals and by meal, reported as the difference between the app's estimate and the reference weight, macronutrient content, or energy content.

Meal characteristics	Absolute error		Bias		Limits of agreement, g
	Absolute value, g, mean (SD)	Relative value, %, mean (SD)	Absolute value, g, mean (SD)	Relative value, %, mean (SD)	
Weight					
All meals	35.1 (42.8)	14.0 (12.2)	19.3 (52.1)	5.4 (17.8)	–84.8, 123.4
Breakfast	30.9 (34.6)	9.5 (7.5)	5.2 (46.8)	0.2 (12.3)	–88.3, 98.7
Cooked meals	62.9 (53.8)	18.2 (14.7)	53.7 (63.6)	15.2 (18.0)	–73.4, 180.8
Snacks	11.5 (14.6)	14.2 (12.4)	–1.0 (18.7)	1.0 (19.2)	–38.5, 36.4
Carbohydrate					
All meals	5.5 (5.1)	14.8 (10.9)	1.0 (7.5)	2.9 (18.3)	–13.9, 15.9
Breakfast	7.1 (5.5)	12.1 (7.9)	–0.5 (9.1)	–1.0 (14.7)	–18.7, 17.7
Cooked meals	6.5 (5.2)	18.2 (11.6)	3.2 (7.8)	8.6 (20.3)	–12.4, 18.7
Snacks	2.9 (3.7)	14.2 (12.5)	0.3 (4.8)	1.0 (19.2)	–9.3, 9.8
Protein					
All meals	2.4 (5.6)	13.0 (13.8)	1.7 (5.9)	5.6 (15.2)	–10.0, 13.4
Breakfast	1.0 (1.1)	7.3 (4.6)	0.0 (1.5)	0.0 (8.8)	–3.1, 3.1
Cooked meals	5.6 (8.9)	17.4 (18.9)	5.3 (9.1)	15.3 (20.8)	–12.9, 23.5
Snacks	0.5 (0.8)	14.2 (12.7)	–0.3 (0.9)	1.4 (19.4)	–2.0, 1.5
Fat					
All meals	1.3 (1.7)	12.3 (12.8)	0.5 (2.1)	5.7 (16.9)	–3.8, 4.7
Breakfast	1.2 (1.3)	8.4 (8.3)	0.4 (1.8)	3.1 (11.6)	–3.1, 3.9
Cooked meals	1.6 (2.3)	14.4 (16.0)	1.3 (2.4)	12.1 (18.0)	–3.6, 6.2
Snacks	1.1 (1.6)	14.0 (12.8)	–0.4 (1.9)	1.8 (19.2)	–4.1, 3.5
Energy					
All meals	41.2 (42.5)[a]	12.7 (10.8)	15.5 (57.4)[a]	4.1 (16.2)	–99.4, 130.3[a]
Breakfast	40.4 (30.5)[a]	9.2 (6.2)	2.1 (51.6)[a]	0.4 (11.4)	–101.1, 105.4[a]
Cooked meals	59.1 (58.0)[a]	14.7 (12.3)	47.8 (68.2)[a]	11.0 (15.8)	–88.7, 184.2[a]
Snacks	24.1 (26.9)[a]	14.2 (12.4)	–3.5 (36.5)[a]	1.0 (19.2)	–76.4, 69.5[a]

[a]kcal.

Figure 2. Bland-Altman plot illustrating the difference between the estimated and reference meal weights.



The mean (SD) absolute error of the estimated carbohydrate content for all meals was 5.5 g (5.1 g; 14.8% [10.9%]), and the mean (SD) bias was 1.0 g (7.5 g; 2.9% [18.3%]). The 95% limits of agreement were –13.9 g and 15.9 g (Figure 3). The mean (SD) absolute error of the estimated protein content for all meals was 2.4 g (5.6 g; 13.0% [13.8%]), and the mean (SD) bias was 1.7 g (5.9 g; 5.6% [18.2%]). The 95% limits of agreement were –10.0 g and 13.4 g (Figure 4). The mean (SD) absolute error of the estimated fat content for all meals was 1.3 g (1.7 g; 12.3% [12.8%]), and the mean (SD) bias was 0.5 g (2.1 g; 5.7% [16.9%]). The 95% limits of agreement were –3.8 g and 4.7 g (Figure 5). The mean (SD) absolute error of the estimated energy content for all meals was 41.2 kcal (42.5 kcal; 12.7% [10.8%]), and the mean (SD) bias was 15.5 kcal (57.4 kcal; 4.1% [16.2%]). The 95% limits of agreement were –99.4 kcal and 130.3 kcal (Figure 6).

Figure 3. Bland-Altman plot illustrating the difference between the estimated and reference carbohydrate content of the meals.

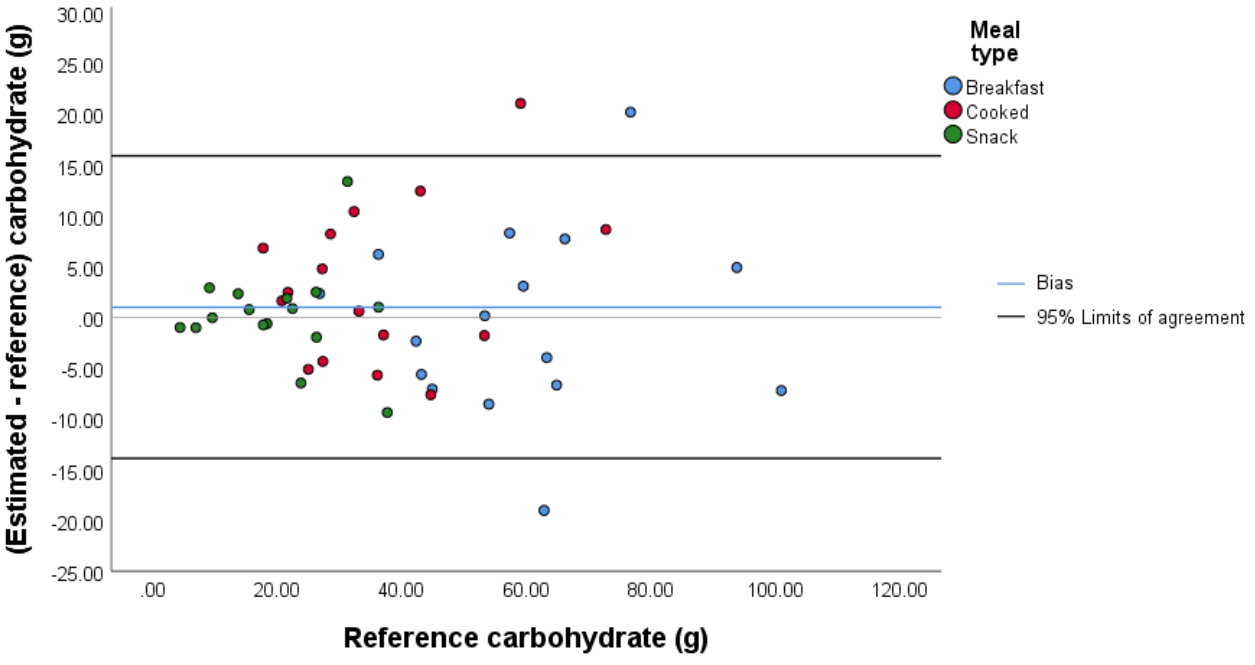


Figure 4. Bland-Altman plot illustrating the difference between the estimated and reference protein content of the meals.

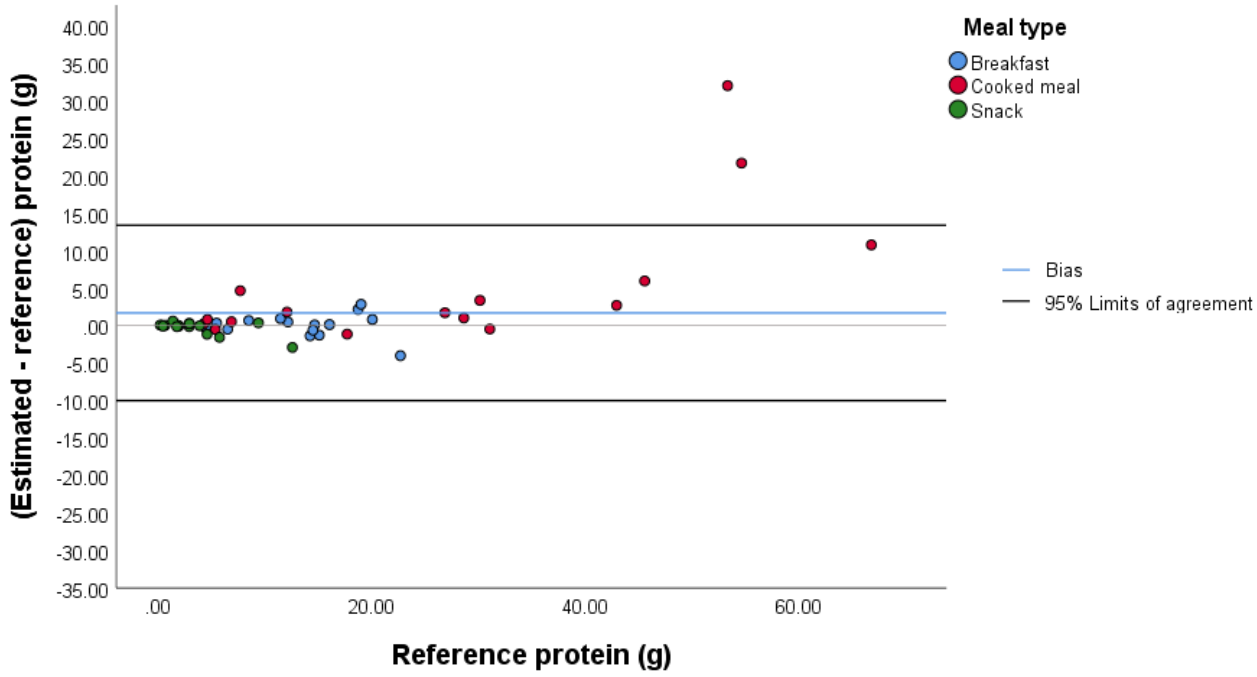


Figure 5. Bland-Altman plot illustrating the difference between the estimated and reference fat content of the meals.

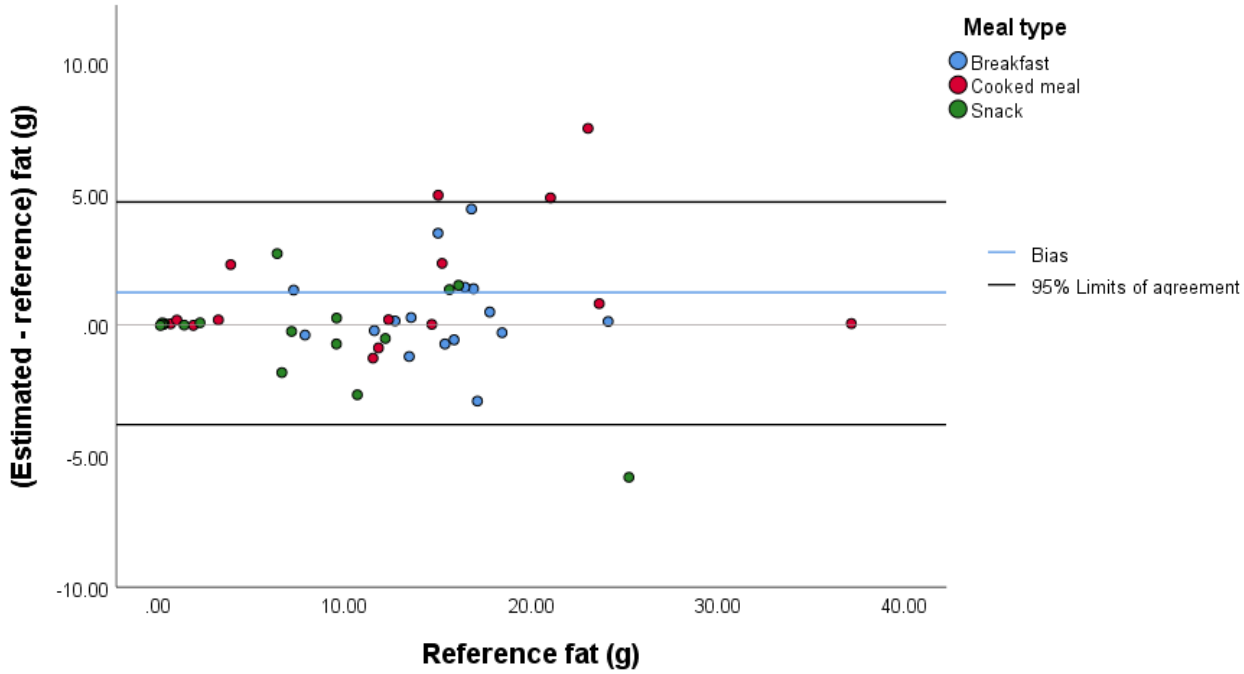
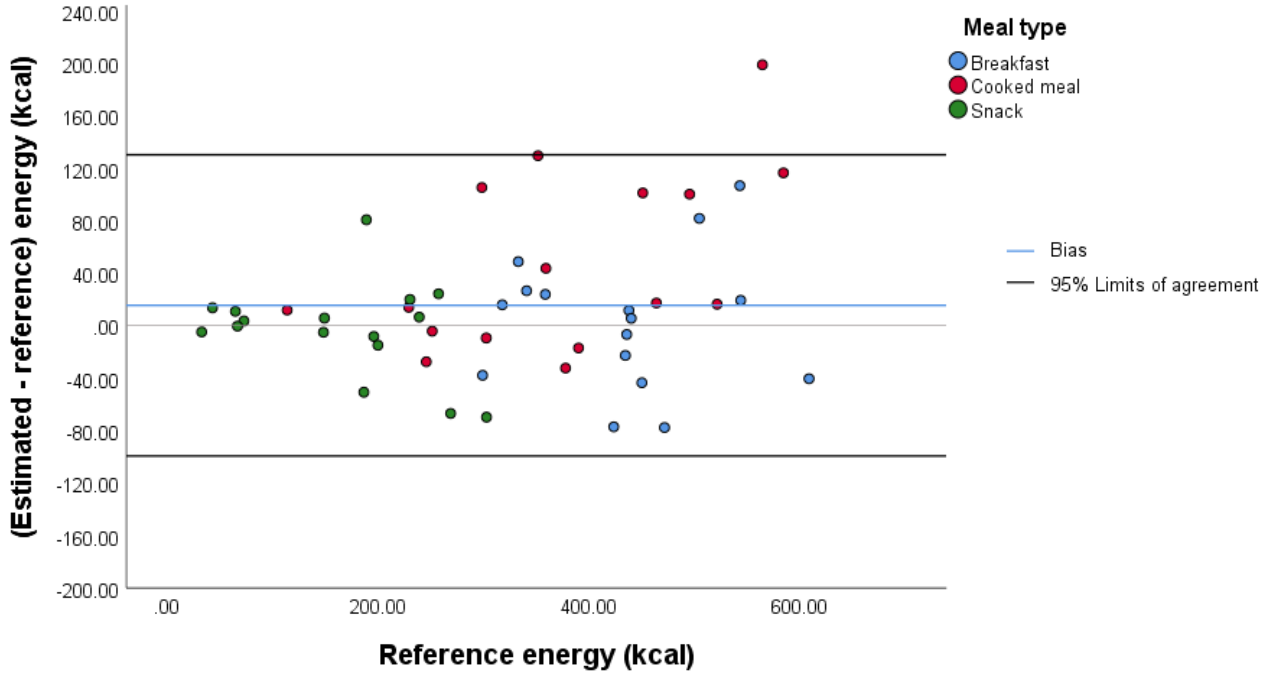


Figure 6. Bland-Altman plot illustrating the difference between the estimated and reference energy content of the meals.



While viewing angle had no significant influence on the accuracy of macronutrient and energy estimation (*P*=.96, *P*=.83, *P*=.99, *P*=.73, and *P*=.70 for absolute errors in weight, carbohydrate, protein, fat, and energy content, respectively), we observed a significant effect of meal type on the accuracy for all macronutrients and energy content (*P*=.001, *P*=.001, *P*=.001, *P*=.002, and *P*=.005 for absolute errors in weight, carbohydrate, protein, fat, and energy content, respectively). The mean bias values of the cooked meals for carbohydrate, protein, and fat content were significantly higher than for the breakfasts, with a mean (SD) difference in bias of 18.3% (4.9%) for carbohydrate, 19.6% (4.9%) for protein, and 17.7% (5.1%) for fat (all *P*<.001). The comparison of bias for snacks relative to the bias of cooked meals resulted in marginal outcomes for fat (*P*=.02), carbohydrate (*P*=.08), and protein (*P*=.07). Furthermore, the bias for snacks was not different from the bias for breakfasts for all macronutrients (*P*=.37, *P*=.26, *P*=.23, *P*=.79, and *P*=.50 for weight, carbohydrate, protein, fat, and energy content, respectively).

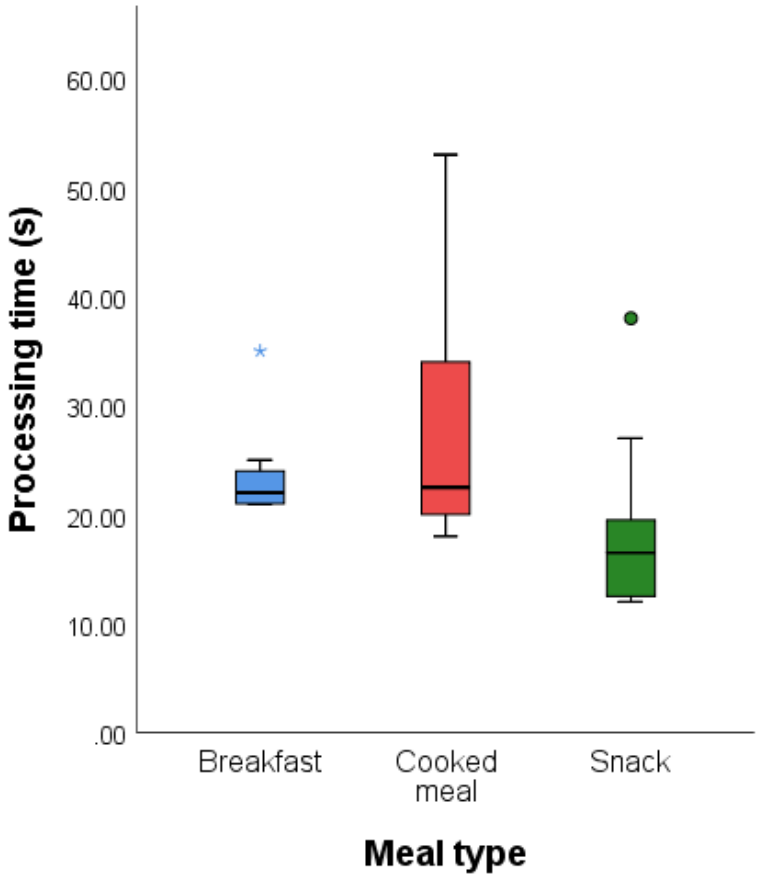
Segmentation Performance

In 7 of the 128 items (5.5%), segmentation required manual adjustment. The intersection over union of unadjusted to adjusted segmentation area was 71.8%.

Processing Time

Mean (SD) processing time across all meal types was 22.9 seconds (8.6 seconds). Processing time was significantly lower for snacks (mean 17.9 seconds, SD 7.0 seconds) compared with cooked meals (mean 27.8 seconds, SD 10.8 seconds; mean difference –9.9 seconds, SD 2.7 seconds; *P*<.001). Processing time was lower for breakfast (mean 23.1 seconds, SD 3.5 seconds) compared with cooked meals (mean difference –4.7 seconds, SD 2.8 seconds; *P*=.12). Figure 7 provides the processing time stratified by meal type.

Figure 7. Box-plot of the processing time according to meal type. Box-plots show median values (solid line), interquartile range (IQR; box outline), spread of data points without outliers (whiskers) and outliers identified as 1.5*IQR (symbols).



Discussion

This study evaluated the accuracy of a novel smartphone app that combines depth sensing with computer vision using volumetry to quantify the macronutrient content of meals in a real-life setting. The main findings were as follows: the accuracy was adequate across all macronutrients, the accuracy differed according to meal type (lower for cooked meals than for snacks and breakfast), segmentation was good, and processing was fast.

When compared with previous reports of apps using computer vision without depth sensors, the present app had comparable, or even superior, accuracy. Rhyner et al [16] reported a mean absolute error of 26.2% in carbohydrate content when assessing 60 cooked meals with non-overlapping food items. In a further preclinical study assessing the accuracy of the prototype used by Rhyner et al and based on 54 cooked meals, the mean absolute error in carbohydrate quantification was 14.8 g, which corresponds to 24.7% for a meal carbohydrate content of 60 g [17]. In contract, with the app in the present study, mean absolute errors in macronutrient content estimation ranged from 12.3% (fat) to 15% (carbohydrate).

Of note, two recent studies assessing the accuracy of image–based food quantification using volume as a reference metric reported mean absolute errors in volume estimation of 7.2% [12] and 5.8% [18] based on the assessment of 5 and 20 food items, respectively. These slightly smaller errors compared to those in this study can be explained by the different reference metric used to define the system accuracy (error in estimated volume versus error in estimated weight and consequently macronutrient content). Of note, errors in weight estimation have two potential sources: inaccuracies in volume and density estimation. Additionally, operational aspects of the previously reported systems differ from those in this study. Xu et al [12] used a complex multi-step approach including reference objects, while Makhsous et al [13] added a depth sensor with structured light to the smartphone and complemented their approach with video sequences, significantly increasing the complexity of the workup. These differences highlight the important tradeoffs between accuracy and usability.

Of note, this study revealed a comparably short processing time, ranging from 18 seconds for snacks to 29 seconds for cooked meals. This is faster than those reported in previous studies, where processing times generally exceeded the limit of 1 minute [19]. This highlights the usability of the present system even when applied to meals in a real-life setting.

The accuracy of the tested app differed according to meal type and was lower for cooked meals than for breakfasts and snacks. This might have resulted from the different levels of complexity in terms of scene analysis of the respective meals. Whereas the breakfasts and snacks had food items that were clearly separated from each other, the cooked meals had food items with touching

borders or a certain degree of overlap. Notably, the angle of image capture did not affect the estimation accuracy in this study, indicating the flexibility, usability, and robustness of the system.

We acknowledge a number of limitations of this study. First, the assessment was limited to meals provided by the hospital kitchen, preventing a generalized statement on the accuracy. However, the system was tested using real-life meals, underscoring its potential use in practice. Second, the system was limited to a single type of smartphone (iPhone X) with a depth sensor, precluding statements on the performance of the software combined with different hardware components. However, this approach supports the strength of providing a commercially available tool. Third, the depth sensor limited the reconstruction to $1/20^{th}$ of the resolution and with lower depth precision than with a passive depth sensor (dual camera approach). However, the use of a depth sensor foregoes the need for fiducial markers, rendering it more convenient to users. Fourth, we served all meals on one plate or bowl type, possibly reducing the variation in volume estimation that was unrelated to the depth sensor. Finally, this study exclusively focused on the accuracy of volume quantification and did not consider food recognition.

When considering both observed accuracy and usability of the present system, the field of potential use appears broad. Such a system may be of interest in the medical sector to assist with nutritional counseling and management of patients with metabolic disorders (eg, diabetes mellitus, obesity) or at risk of malnutrition. Beyond this, such a system may be valuable in nutritional epidemiology due to the potential to systematically and accurately monitor dietary intake on a large scale.

In conclusion, this study evaluated the accuracy of a novel smartphone app with integrated depth sensing and found a high level of accuracy in volumetric macronutrient and energy estimation across a broad set of meals in a real-life setting. In addition, the system demonstrated high segmentation performance and low processing time, highlighting its usability.

Acknowledgments

We thank the staff from the Bern University Hospital Kitchen (Thomas Walser, Beat Blum, Vinzenz Meier) for provision of food items and tableware. We also extend special thanks to Michèle Monnard for support with data management and to the developer of the software SNAQ (Aurelian Briner, Nico Previtali, Lukas Frischknecht, and the team) for the pleasant collaboration. This project was supported by the Diabetes Center Bern and the Berner Kantonalbank (BEKB).

Authors' Contributions

DH, CN, CS, and LB conceived and designed the study. DH, CK, CL, and RJ acquired the data. DH, CN, and LB analyzed the data. DH, CN, JD, and LB interpreted the data. All authors were involved in drafting or revising the article for important intellectual content. All authors approved the final version of the manuscript. CS and LB are responsible for the integrity of the work as a whole.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Food items assessed using the SNAQ smartphone app.
[DOCX File , 15 KB - mhealth_v8i3e15294_app1.docx]

Multimedia Appendix 2
Examples of good and bad segmentation by the SNAQ app during test meal analysis.
[PNG File , 1766 KB - mhealth_v8i3e15294_app2.png]

Multimedia Appendix 3
Mean macronutrient and energy content of the tested meals.
[DOCX File , 16 KB - mhealth_v8i3e15294_app3.docx]

Multimedia Appendix 4
Examples of the assessed meal types.
[PNG File , 1301 KB - mhealth_v8i3e15294_app4.png]

Multimedia Appendix 5
Estimation accuracy, as determined using the error metrics of the analyzed meals.
[DOCX File , 15 KB - mhealth_v8i3e15294_app5.docx]

References

1. American Diabetes Association. Evidence-based nutrition principles and recommendations for the treatment and prevention of diabetes and related complications. Diabetes Care 2002 Jan;25(1):202-212. [doi: 10.2337/diacare.25.1.202] [Medline: 11772917]
2. Raynor HA, Champagne CM. Position of the Academy of Nutrition and Dietetics: Interventions for the Treatment of Overweight and Obesity in Adults. J Acad Nutr Diet 2016 Jan;116(1):129-147. [doi: 10.1016/j.jand.2015.10.031] [Medline: 26718656]
3. Livingstone MB, Prentice AM, Strain JJ, Coward WA, Black AE, Barker ME, et al. Accuracy of weighed dietary records in studies of diet and health. BMJ 1990 Mar 17;300(6726):708-712 [FREE Full text] [doi: 10.1136/bmj.300.6726.708] [Medline: 2386561]
4. Ordabayeva N, Chandon P. In the eye of the beholder: Visual biases in package and portion size perceptions. Appetite 2016 Aug 01;103:450-457. [doi: 10.1016/j.appet.2015.10.014] [Medline: 26482283]
5. Beasley J, Riley WT, Jean-Mary J. Accuracy of a PDA-based dietary assessment program. Nutrition 2005 Jun;21(6):672-677. [doi: 10.1016/j.nut.2004.11.006] [Medline: 15925290]
6. Maringer M, Van't Veer P, Klepacz N, Verain MCD, Normann A, Ekman S, et al. User-documented food consumption data from publicly available apps: an analysis of opportunities and challenges for nutrition research. Nutr J 2018 Jun 09;17(1):59 [FREE Full text] [doi: 10.1186/s12937-018-0366-6] [Medline: 29885653]
7. Chowdhary CL. 3D Object Recognition System Based On Local Shape Descriptors and Depth Data Analysis. CSENG 2019 Jan 10;12(1):18-24. [doi: 10.2174/2213275911666180821092033]
8. Ciocca G, Napoletano P, Schettini R. Food Recognition: A New Dataset, Experiments, and Results. IEEE J. Biomed. Health Inform 2017 May;21(3):588-598. [doi: 10.1109/jbhi.2016.2636441]
9. Anthimopoulos M, Dehais J, Shevchik S, Ransford BH, Duke D, Diem P, et al. Computer vision-based carbohydrate estimation for type 1 patients with diabetes using smartphones. J Diabetes Sci Technol 2015 May;9(3):507-515 [FREE Full text] [doi: 10.1177/1932296815580159] [Medline: 25883163]
10. He H, Kong F, Tan J. DietCam: Multiview Food Recognition Using a Multikernel SVM. IEEE J. Biomed. Health Inform 2016 May;20(3):848-855. [doi: 10.1109/jbhi.2015.2419251]
11. Chen H, Jia W, Yue Y, Li Z, Sun Y, Fernstrom JD, et al. Model-based measurement of food portion size for image-based dietary assessment using 3D/2D registration. Meas Sci Technol 2013 Oct;24(10) [FREE Full text] [doi: 10.1088/0957-0233/24/10/105701] [Medline: 24223474]
12. Xu C, He Y, Khanna N, Boushey C, Delp E. Model-Based Food Volume Estimation Using 3D Pose. In: Proc Int Conf Image Proc. 2013 Sep Presented at: International Conference on Image Processing; 2013; Melbourne, Australia p. 2534-2538 URL: http://europepmc.org/abstract/MED/28572746 [doi: 10.1109/ICIP.2013.6738522]
13. Makhsous S, Mohammad HM, Schenk JM, Mamishev AV, Kristal AR. A Novel Mobile Structured Light System in Food 3D Reconstruction and Volume Estimation. Sensors (Basel) 2019 Jan 29;19(3) [FREE Full text] [doi: 10.3390/s19030564] [Medline: 30700041]
14. Derpanis K. Overview of the RANSAC Algorithm. Computer Science, York University, Tech. Rep 2010.
15. Heise T, Meiffren G, Alluis B, Seroussi C, Ranson A, Arrubla J, et al. BioChaperone Lispro versus faster aspart and insulin aspart in patients with type 1 diabetes using continuous subcutaneous insulin infusion: A randomized euglycemic clamp study. Diabetes Obes Metab 2018 Dec 18. [doi: 10.1111/dom.13621] [Medline: 30565407]
16. Rhyner D, Loher H, Dehais J, Anthimopoulos M, Shevchik S, Botwey RH, et al. Carbohydrate Estimation by a Mobile Phone-Based System Versus Self-Estimations of Individuals With Type 1 Diabetes Mellitus: A Comparative Study. J Med Internet Res 2016 May 11;18(5):e101 [FREE Full text] [doi: 10.2196/jmir.5567] [Medline: 27170498]
17. Vasiloglou M, Mougiakakou S, Aubry E, Bokelmann A, Fricker R, Gomes F, et al. A Comparative Study on Carbohydrate Estimation: GoCARB vs. Dietitians. Nutrients 2018 Jun 07;10(6) [FREE Full text] [doi: 10.3390/nu10060741] [Medline: 29880772]
18. Puri MZ, Yu Q, Divakaran A, Sawhney H. Recognition and volume estimation of food intake using a mobile device. In: IEEE.: Workshop on Application of Computer Vision; 2009 Presented at: Workshop on Applications of Computer Vision (WACV); 7-8 December, 2009; Snowbird, UT. [doi: 10.1109/wacv.2009.5403087]
19. Eldridge A, Piernas C, Illner A, Gibney M, Gurinović M, de Vries J, et al. Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients 2018 Dec 28;11(1) [FREE Full text] [doi: 10.3390/nu11010055] [Medline: 30597864]

Abbreviations

IQR: interquartile range.

Edited by G Eysenbach; submitted 28.06.19; peer-reviewed by F Zhu, N Alves, S Guness; comments to author 14.10.19; revised version received 05.11.19; accepted 16.12.19; published 25.03.20.

Please cite as:
Herzig D, Nakas CT, Stalder J, Kosinski C, Laesser C, Dehais J, Jaeggi R, Leichtle AB, Dahlweid FM, Stettler C, Bally L
Volumetric Food Quantification Using Computer Vision on a Depth-Sensing Smartphone: Preclinical Study
JMIR Mhealth Uhealth 2020;8(3):e15294
URL: http://mhealth.jmir.org/2020/3/e15294/
doi:10.2196/15294
PMID:32209531

Original Paper

Quantitative Screening of Cervical Cancers for Low-Resource Settings: Pilot Study of Smartphone-Based Endoscopic Visual Inspection After Acetic Acid Using Machine Learning Techniques

Jung Kweon Bae[1*], MEng; Hyun-Jin Roh[2*], MD; Joon S You[1], PhD; Kyungbin Kim[3], MD; Yujin Ahn[1], MEng; Sanzhar Askaruly[1], MEng; Kibeom Park[1], PhD; Hyunmo Yang[1], PhD; Gil-Jin Jang[4], PhD; Kyung Hyun Moon[5], MD; Woonggyu Jung[1], PhD

[1]Department of Biomedical Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea

[2]Department of Obstetrics and Gynaecology, University of Ulsan College of Medicine, Ulsan University Hospital, Ulsan, Republic of Korea

[3]Department of Pathology, Ulsan University Hospital, Ulsan, Republic of Korea

[4]School of Electronics Engineering, Kyungpook National University, Daegu, Republic of Korea

[5]Department of Urology, University of Ulsan College of Medicine, Ulsan University Hospital, Ulsan, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Woonggyu Jung, PhD
Department of Biomedical Engineering
Ulsan National Institute of Science and Technology
Rm#709, 110 Bldg
50 Unist-gil
Ulsan, 44919
Republic of Korea
Phone: 82 1084640110
Email: wgjung@unist.ac.kr

Abstract

Background: Approximately 90% of global cervical cancer (CC) is mostly found in low- and middle-income countries. In most cases, CC can be detected early through routine screening programs, including a cytology-based test. However, it is logistically difficult to offer this program in low-resource settings due to limited resources and infrastructure, and few trained experts. A visual inspection following the application of acetic acid (VIA) has been widely promoted and is routinely recommended as a viable form of CC screening in resource-constrained countries. Digital images of the cervix have been acquired during VIA procedure with better quality assurance and visualization, leading to higher diagnostic accuracy and reduction of the variability of detection rate. However, a colposcope is bulky, expensive, electricity-dependent, and needs routine maintenance, and to confirm the grade of abnormality through its images, a specialist must be present. Recently, smartphone-based imaging systems have made a significant impact on the practice of medicine by offering a cost-effective, rapid, and noninvasive method of evaluation. Furthermore, computer-aided analyses, including image processing–based methods and machine learning techniques, have also shown great potential for a high impact on medicinal evaluations.

Objective: In this study, we demonstrate a new quantitative CC screening technique and implement a machine learning algorithm for smartphone-based endoscopic VIA. We also evaluated the diagnostic performance and practicability of the approach based on the results compared to the gold standard and from physicians' interpretation.

Methods: A smartphone-based endoscope system was developed and applied to the VIA screening. A total of 20 patients were recruited for this study to evaluate the system. Overall, five were healthy, and 15 were patients who had shown a low to high grade of cervical intraepithelial neoplasia (CIN) from both colposcopy and cytology tests. Endoscopic VIA images were obtained before a loop electrosurgical excision procedure for patients with abnormal tissues, and their histology tissues were collected. Endoscopic VIA images were assessed by four expert physicians relative to the gold standard of histopathology. Also, VIA features were extracted from multiple steps of image processing techniques to find the differences between abnormal (CIN2+) and normal (≤CIN1). By using the extracted features, the performance of different machine learning classifiers, such as k-nearest neighbors (KNN), support vector machine, and decision tree (DT), were compared to find the best algorithm for VIA. After determining the best performing classifying model, it was used to evaluate the screening performance of VIA.

Results: An average accuracy of 78%, with a Cohen kappa of 0.571, was observed for the evaluation of the system by four physicians. Through image processing, 240 sliced images were obtained from the cervicogram at each clock position, and five features of VIA were extracted. Among the three models, KNN showed the best performance for finding VIA within holdout 10-fold cross-validation, with an accuracy of 78.3%, area under the curve of 0.807, a specificity of 80.3%, and a sensitivity of 75.0%, respectively. The trained model performed using an unprovided data set resulted in an accuracy of 80.8%, specificity of 84.1%, and sensitivity of 71.9%. Predictions were visualized with intuitive color labels, indicating the normal/abnormal tissue using a circular clock-type segmentation. Calculating the overlapped abnormal tissues between the gold standard and predicted value, the KNN model overperformed the average assessments of physicians for finding VIA.

Conclusions: We explored the potential of the smartphone-based endoscopic VIA as an evaluation technique and used the cervicogram to evaluate normal/abnormal tissue using machine learning techniques. The results of this study demonstrate its potential as a screening tool in low-resource settings.

(JMIR Mhealth Uhealth 2020;8(3):e16467) doi:10.2196/16467

KEYWORDS

smartphone-based endoscope; smartphone VIA; machine learning; cervical cancer screening; low-resource settings

Introduction

According to the International Agency for Research on Cancer and GLOBOCAN 2018, cervical cancer (CC) is the fourth most frequent cancer in women worldwide [1], and approximately 90% of the global cervical cancer deaths in 2015 occurred in low- and middle-income countries [2,3]. Although CC is regarded as a highly preventable and curable cancer, it is still one of the leading causes of mortality in low-resource settings and developing countries due to their lack of sustainable screening programs and limited infrastructure [3-5]. CC can be readily managed when it is found in the precancerous stages through routine screening methods, such as a cytology-based test. The most popular and affordable method for CC screening in low-resource countries is the use of visual inspection with acetic acid (VIA). Since VIA offers relatively simple, cost-effective visual feedback, it can even provide treatment on the same day of a screening visit [4-7]. In VIA, the topical application of 4-5% acetic acid to the cervix transforms abnormal squamous epithelium to a dense white color, while normal epithelium presents as a light pink color. Despite its simplicity, VIA provides sufficient sensitivity and specificity to identify the cancerous lesion; thus, it has been widely promoted and recommended as an alternative to the conventional cytology test (ie, the Pap smear) [4-8]. Nonetheless, visual inspection methods have been found to be subjective and the range of diagnostic performance varies widely, with significantly better results obtained by physicians than by nurses [8]. Unfortunately, in many developing countries, trained physicians who can interpret VIA correctly may not be readily available [5,7,8].

Digital images of the cervix after application of acetic acid, or digital cervicography, have been significantly important for improving quality control. It is a very efficient way of minimizing interpreters' subjectivity by capturing higher resolution images for post-screening analysis [9-13]. Moreover, digital images can be transmitted or shared through the internet with long-distance experts, thus closing the gap in human resources [13]. Recent advances in smartphone technologies have opened new possibilities for cervical screening in low-resource settings, thus overcoming the limitations of colposcopy, including the device's bulkiness, high-cost, electricity dependency, and constant maintenance need [14-20].

The smartphone is a highly integrated platform that includes various functionalities, easy accessibility, a user-friendly interface, ubiquitous internet, and communication technologies [21]. The high-definition camera in a smartphone has especially made an impact on the practice of medicine by offering cost-effective, rapid, and noninvasive imaging capabilities [21-25]. Smartphone-based cervical screening has been proven feasible and validated for quality assurance in low resource settings [14-20]. Smartphone-based digital visual inspection following application of acetic acid has been demonstrated for higher diagnostic accuracy and reduction of the variability of detection rate. Although digital images are very effective in various ways [5], implementation of remote expert consultation is still challenging due to the lack of reliable broadband connections in remote areas [19].

On the other hand, automated interpretation of data and classification of cervical images for instant diagnostic conclusions will enable on-site treatments to be delivered without delays [26-31]. To date, various image processing and interpretation methods have been successfully applied to VIA using such features as aceto-whitening, blood vessel formation, and texture of the surface [29-31]. Previous works have shown that automated classification of VIA can perform as well as experts' qualitative assessment of colposcopic images [26-28]. Also, auxiliary processing methods, such as elimination of speculum reflection and determination of the region of interest (ROI), can further improve the overall performance of image processing outcomes [31]. Automated quantification of VIA based on modern image processing and machine learning techniques could be a very promising platform for cervical screening in low-resource settings. However, a fully automated diagnostic performance using smartphone-based cervical images has not been introduced, despite a clear need and potential.

In this study, we demonstrate a new quantitative CC screening technique by implementing a machine learning algorithm for smartphone-based endoscopic VIA. Our method can provide digital images as well as an automated diagnostic classification for comprehensive and intuitive feedback to a clinician. We

have evaluated the diagnostic performance of the system through quantitative comparison to the gold standard of cytology and physicians' interpretation of the digital images. This approach would extend cervical cancer screening to remote populations who do not have access to experienced colposcopists.

Methods

Smartphone-Based Endoscope System

We developed a miniaturized endoscope system by assembling an endoscopic probe and smartphone with customized ancillary components, as reported previously [25]. The smartphone-based endoscope system and its components are illustrated in Figure 1. The system is composed of 3 major components: (1) Customized coupler for universal attachment of endoscopic probes, generated by three-dimensional (3D) printer (Stratasys, Objet260 Connex2); (2) smartphone case, also generated by 3D printer; and (3) an optical adapter used for magnification, which is placed between the endoscopic probe and the smartphone camera, as shown in Figure 1, part (a). Incorporating achromatic and aspherized achromatic lenses that had 40-millimeter and 14-millimeter focal lengths (Figure 1, part [b]), respectively, we obtained approximately 4× optical magnification. The image can be further magnified up to approximately 12× with a smartphone's digital zoom feature. A portable light source was attached to the illumination port of an endoscopic probe. Images were acquired with the home-built android application that features control functions such as compensation of the rotated images caused by the lens, camera controls, including zoom, ISO sensitivity, white balance, resolution size, and exposure adjustments, and the option to save files.

Optical elements of the adapter significantly improved the cervix image captured with a smartphone camera alone. Images of the central part of the Unites States air force (USAF) resolution target were captured with/without a smartphone endoscope (Figure 1, part [c]). The device was placed at 150 millimeters and 300 millimeters away from the target, where prior smartphone-based VIA [14,19] and routine colposcopy [32] are conducted. For the endoscope system, we placed the distal end of the probe at 20 millimeters away from the target, where the whole ectocervix was well defined in the field of view. As shown in Figure 1, part (d), smartphone-based endoscopy achieved the best resolution from a line plot representing Group 2, Element 4, from part (c), in the resolution target.

Figure 1. Schematic of the system. (a) 3D modeling of the smartphone-based endoscope system. (b) Optics of the customized zoom lens. (c) Images of the resolution target taken without any optics adapter and our system, respectively. (d) Group 2, element 4 from each image was described as a line plot. 3D: three-dimensional; AU: arbitrary unit.

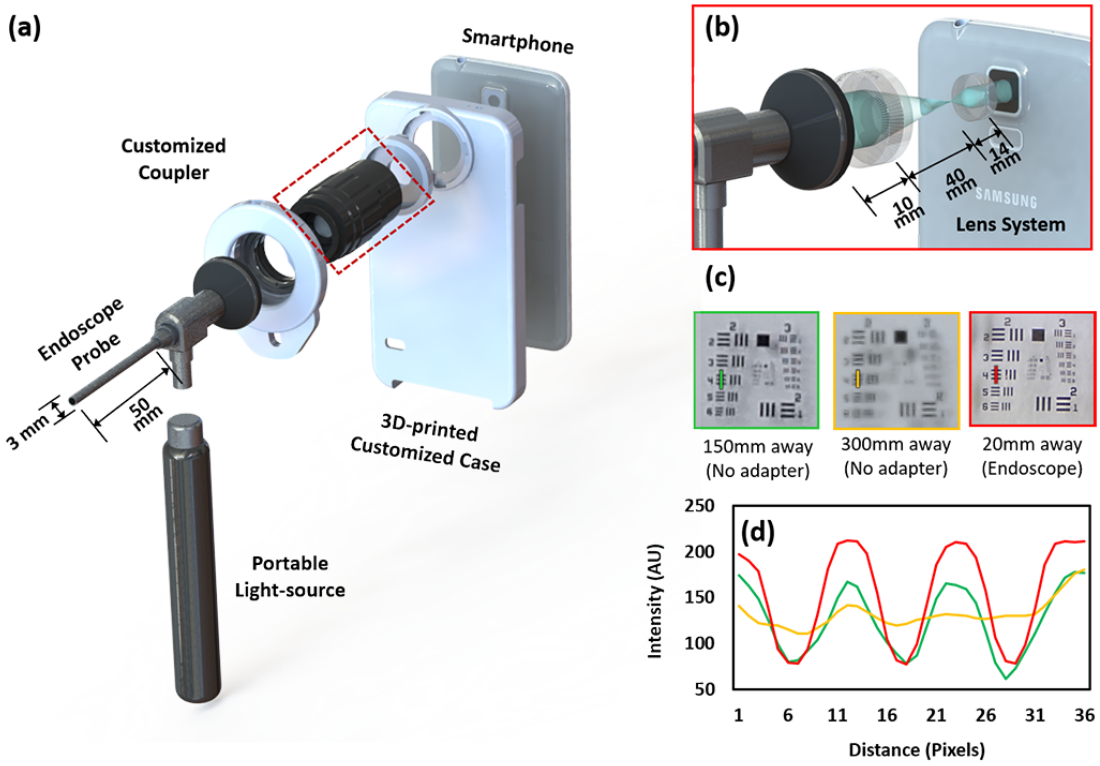


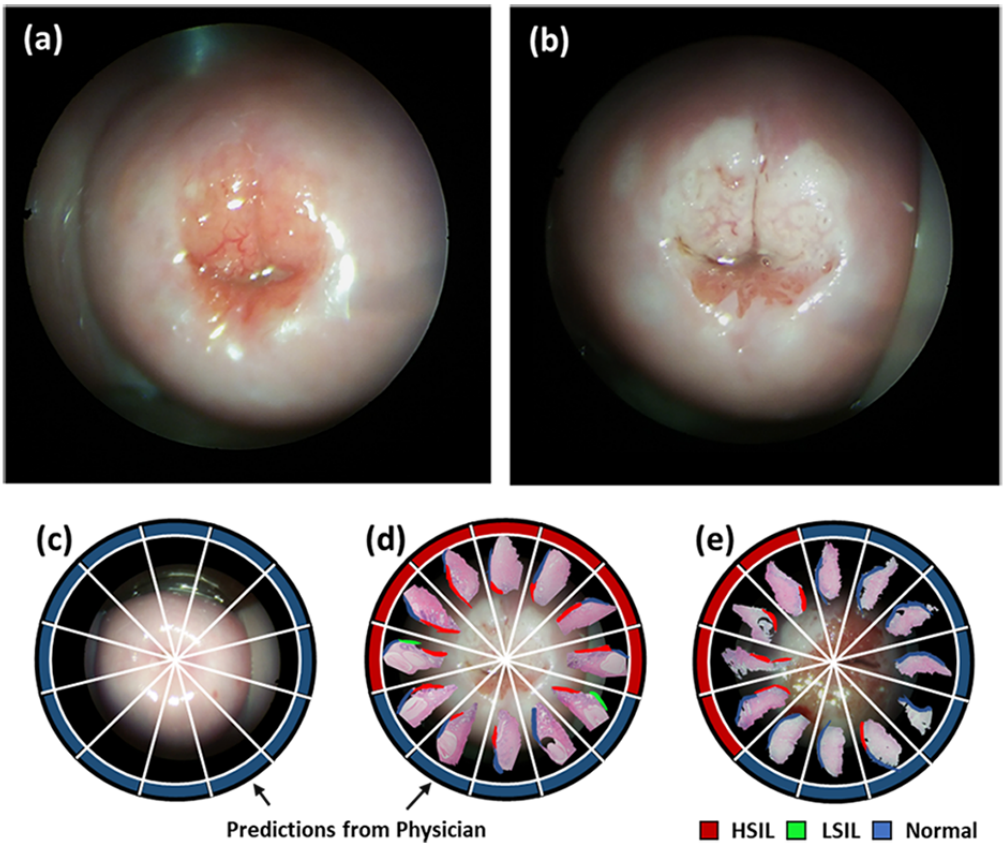
Image Acquisition From Clinic

Following a protocol approved by the Ulsan University Hospital Institutional Review Board, we collected smartphone-based VIA images using an endoscope in human subjects. In Ulsan University Hospital, VIA was routinely performed to visualize the margin of the suspicious tissue before loop electrosurgical excision procedure (LEEP). Therefore, each imaging session was conducted before LEEP, in an operating room. A rigid endoscopic probe (Medstar, Otoscope, 0°, Ø4, and 50 mm length) was inserted inside a subject's vagina where speculum had already been placed. Since the endoscope is thin, smartphone-based endoscopic VIA imaging was performed in a noninvasive and noncontact manner. All patients who underwent LEEP had already been potential candidates to have cervical intraepithelial neoplasia (CIN2+), which was

determined by previous cytology-based tests and colposcopy. A typical procedure in this study took less than five minutes without causing undue burden on volunteers and delaying the treatment. First, one minute was used to take images before the application of the acetic acid, then the next 1-2 minutes was used to apply the 3-5% acetic acid, then another minute for waiting, and then the last minute was used to take VIA images. Figure 2, parts (a) and (b), show representative images of the smartphone-based endoscopic VIA. For patients who underwent LEEP, 12 tissue sections were collected at each clock position from the excised ectocervix. For this study, physicians labeled the CIN grades in colors, as shown Figure 2, parts (c)-(e). A total of 20 patients aged 20 years old or older participated in this study. Among them, five volunteers were normal (CIN1-), and 15 were confirmed abnormal (CIN1+) using the gold standard cytology test. Normal cervix status of the five subjects was verified as such by cytology test and colposcopy, so no LEEP was performed and no tissues were collected from them.

All captured images, including before and after application of the acetic acid, were sent to expert physicians for review. A total of four experts with professional experience, ranging from 12-20 years, participated and were kept blind to the results of the machine learning and the cytology. Physicians' interpretations were based only on VIA features without any additional information given to them. During the image reviews, physicians labeled the directional information of the tissue region that contained suspicious abnormal features (Figure 2). In this study, tissues, including normal and CIN1, were considered to be normal, and CIN2+ was considered to be abnormal because only CIN2+ requires treatment [26].

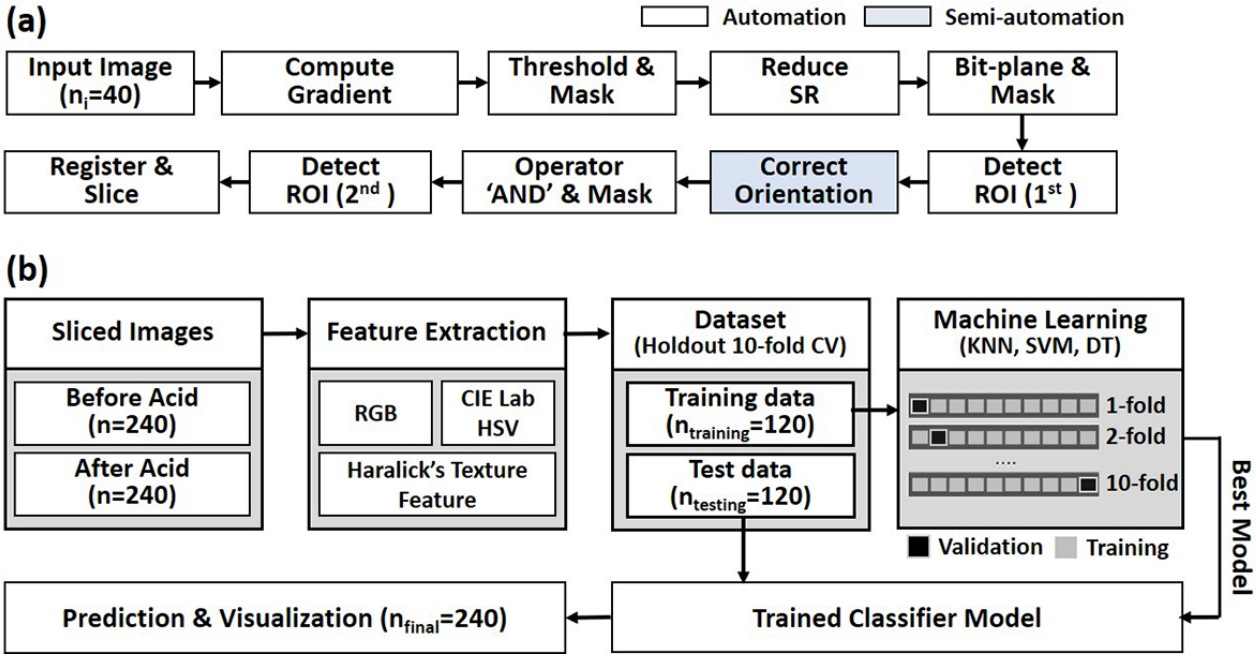
Figure 2. Smartphone-based endoscopic cervicogram. (a) Cervicogram of before acetic acid application. (b) Cervicogram of after acetic acid application. (c) VIA- patient (d-e) VIA+ patients with predictions from best among four physicians. Prediction labeled the precancerous regions with colors at each clock position. VIA: Visual inspection with acetic acid.



Preimage Processing

Images from smartphone-based endoscopic VIA contain unnecessary features, such as vaginal walls, speculum, and specular reflections of light, as commonly found in typical cervicograms [28,31]. As these features may affect overall classification accuracy, we performed multi-step image processing to reduce their influence on data analysis (See Figure 3a).

Figure 3. Block diagram of image processing and classification. (a) Preimage processing method, (b) Feature analysis and classification. SR: speculum reflection; ROI: region of interest; RGB: red, green, blue; CIE: Commission Internationale de l'Eclairage; HSV: hue, saturation, value; CV: cross-validation; KNN: K-nearest neighbor; SVM: supportive vector machine; DT: decision tree.



Next, ROI was defined to segment the major cervix region. The major cervix region in images both before and after the application of acetic acid contains high red-channel values [27]. For ROI detection, the red-channel image was transferred to grayscale and separated into multilevel binary images. Here, the 8-bit grayscale of the red-channel image was sliced into eight planes ranging from the least significant bit, 0, to the most significant bit, 7. Most of the seventh and eighth bit-planes represent the major ectocervix regions, so we used these features to generate the mask for ROI segmentation. Due to the slight differences of the cervix images between pre– and post–acetic acid application, we performed an 'AND' operator to segment the overlapping regions.

Before VIA features extraction and classification, image pairs for pre– and post–acetic acid application had to be properly registered. As shown in Multimedia Appendix 1, we manually provided three points as fiducials on each image to locate the center of the cervix (red dot) and both ends of the cervical os (2 black dots). Utilizing the center point, we correctly registered the center of the cervix for every image pair. We also drew a line penetrating the other two points and found the angle of the line from each image. We made this line horizontal by rotating images with respect to the given center point. Further, we cropped the images into 12 pieces, as sectioned in histology. In this image processing method, we used a total of 20 image pairs acquired from volunteers as initial input (n_i=40) and obtained 240 pieces of images that have directional information for both before (n_{cb}=240) and after (n_{ca}=240) acetic acid application. All detailed procedures with representative images for each step can be found in Multimedia Appendix 1.

Feature Analysis and Selection

After the preprocessing of images, VIA features were analyzed to identify the images containing the suspicious lesions. To extract the abnormal features, we inspected RGB color intensity, values in extended color space, and Haralick's texture features [33]. The application of the acetic acid on squamous epithelial areas coagulates the cellular protein and dehydrates the cytoplasm. Images of VIA– cervix, thus, generally showed light pink or very thin white appearances due to the reflection of light from the underlying stroma. On the other hand, VIA+ tissues that are rich in cellular proteins were presented with thick white features which blocked the colors of the stroma. Due to this reason, VIA+ incorporates larger, thick, white areas in images that would exhibit more green and blue intensities in color space [26,34]. From there, we computed the green-to-red and blue-to-red intensity ratio and found the separation between histogram distributions of pre– and post–acetic acid application images. This approach would properly quantify changes in green and blue intensities independent of device variation and level of illumination. The differences of histogram distribution of average (D_{ave}), green-to-red ($D_{G/R}$), blue-to-red ($D_{B/R}$), and average histogram differences of green-to-red and blue-to-red (D_{ave}) can then be defined as Standalone Equation 1, where I(mode) is intensity level or index at mode in histogram.

The standard deviations of the green and blue channels are another important extracted feature. Aceto-white features with higher green and blue channel intensities also exhibited higher standard deviations, which means more green and blue intensities are widely distributed in the histogram. Different color space was also utilized to find the features of VIA. Commission Internationale de l'Eclairage (CIE)*Lab color space was computed to achieve perceptual uniform color space, which is useful to quantitatively distinguish between the colors of an image. The great advantage of CIE*Lab is that it is independent of device and illumination [26].

In this study, the major color included on the cervix was defined by calculating the average of a* channels from each post-acetic application. Due to abnormal vascular formation, such as mosaicism and punctuation, visualization of the uneven surface of the ectocervix was another contrasting feature. Computation of Haralick's texture feature using gray-level cooccurrence matrix (GLCM) may quantify the spatial variation of gray intensity values related to the texture of an image. GLCM measures the probability distributions of different combinations of pixel values. Utilizing the GLCM, several pieces of statistical information, such as contrast, correlation, energy, and homogeneity, can be derived quantitatively to exhibit the texture of the image [35,36]. Here, the GLCM was calculated at four different angles (0°, 45°, 90°, 135°) with an interpixel distance of 5 for the difference of the S channel (HSV color space) from pre– to post–acetic acid images. Different angles measure the features of interest in every direction. Therefore, all four GLCMs were summed before texture calculation.

In this study, we specifically utilized correlation statistics, which provide the extent of correlation between a pixel to its neighbor pixel over the whole image [35]. The correlation statistics (mean difference=0.248; *P*<.001) exhibited a significant difference in the VIA+/– images, yet other texture statistics had shown a very small difference, down to third and fourth decimal points in mean difference (contrast: *P*=.066; homogeneity: *P*=.308; energy: *P*=.249). Therefore, five VIA features were analyzed and selected as potentially useful for diagnostic classification: (1) Average difference of green/blue-to-red histogram index; (2) SD of green channel from post–acetic acid images; (3) SD of blue channel from post-acetic acid images; (4) average value of a* channel; and (5) correlation values of Haralick's texture features from S channel information.

Classification Training and Validation

Using selected features of VIA as predictors, we interpreted the tissue abnormality of a localized region for classification. Machine learning techniques have been widely used and successfully supervised for VIA classification [25,31]. In this work, we examined and selected an appropriate classifier by analyzing three different classifying methods. We performed holdout k-fold cross-validation (k=10), not only to optimize the hyperparameter to avoid overfitting/underfitting problems, but also to select the best performing model. Thus, we randomized the order of the images and used half ($n_{training}$=120) for training classifiers. The other half of the data, an untrained image set ($n_{testing}$=120), was used as the testing set after the optimization of classification models.

While in training, classifiers are validated using k-fold cross-validation (k=10) with histopathology labels as the ground truth. In this method, data is evenly divided into k subsamples. Other k-1 subsamples are then used as training datasets, and then held-out or excluded subsamples are used for validation. Algorithms were repeated k times, with each of the subsamples only utilized once, as the validating data and performance accuracy were calculated by averaging the results from each k-fold [37]. We designed k-nearest neighbors (KNN) with five neighbors based on Euclidian distance, support vector machine (SVM) with a fourth degree of polynomial kernel function (cost=3; gamma=2.2), and decision tree (DT) with a limit of a maximum of 20 nodes, based on Gini's diversity split criterion. These parameters, or hyperparameters, in each classifying method were optimized through the grid-search technique. All predictors were standardized using their corresponding weighted means and weighted standard deviations [28]. By using a validation result in training, receiver operating characteristic (ROC) curves with area under the curve (AUC), accuracy, sensitivity, and specificity of each trained classifier were computed and compared for selecting the best classifiers. Throughout the classifiers, each image was interpreted to either VIA+ or VIA–. The classification process is illustrated in Figure 3.

Results

Direct Evaluation of Smartphone-Based Endoscopic VIA

To determine the feasibility of the smartphone-based endoscope system for the VIA application, four physicians participated and reviewed the image sets (n=20). In these 20 cases, both clinically normal and low-grade squamous intraepithelial lesions (LSILs) were designated as "normal," and high-grade squamous intraepithelial lesions (HSILs) was designated as "abnormal," resulting in 8 normal and 12 abnormal cases for this study. The diagnostic performance assessed by pathologists is summarized in Table 1. Sensitivity ranged from 33.3-83.3%, and specificity was 100% for four physicians. Accuracy of the assessment ranged from 60.0%, with a Cohen kappa of 0.286 (*P*=.068), to 90.0%, with a Cohen kappa of 0.800 (*P*<.001). Overall, an average accuracy of 78%, with Cohen kappa of 0.571 (*P*<.001), was observed for the evaluation of the system.

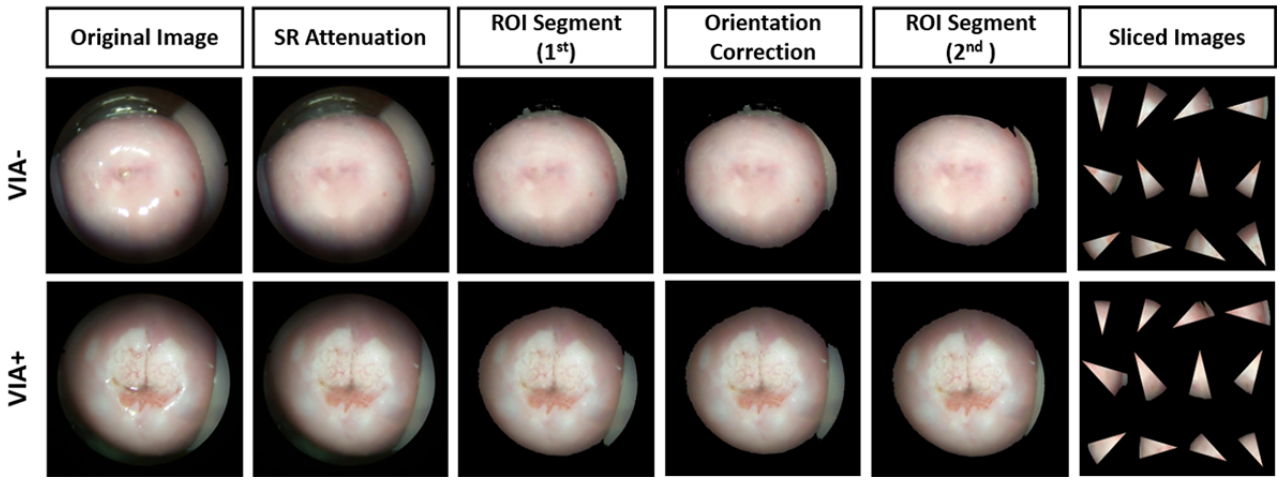
Table 1. Sensitivity, specificity, Cohen kappa value, and *P* value for smartphone-based endoscopic VIA among all observers (n=20).

Physicians	Sensitivity, %	Specificity, %	Accuracy, %	Cohen Kappa	*P* value
Physician 1	75.0	100.0	85.0	0.706	.001
Physician 2	33.3	100.0	60.0	0.286	.068
Physician 3	83.3	100.0	90.0	0.800	<.001
Physician 4	58.3	100.0	75.0	0.528	.007
Average	62.5	100.0	77.5	0.571	<.001

Result of Image Processing

Figure 4 shows the representative images from each image processing step. Specular reflections on the surface of the cervix are removed while preserving visually natural or smooth features. All areas of specular reflection with saturated intensity are correctly localized for all 20 image pairs, including both pre– and post–acetic acid application images. The intensity threshold of 60% for specular reflection removal was effective, as the reflections on the ectocervix regions were identified clearly while minimally affecting general backgrounds. This was feasible because more reflection intensity was obtained in the focal plane compared to that of the unfocused regions. Next, the major cervical region was selected as the ROI and was segmented by applying the seventh and eighth bit-planes of the red channel image as the primary mask. Although most of the artifacts, such as unfocused background and speculum, were successfully eliminated, some images still contained portions of the vaginal wall. These results were likely found when the cervix and vaginal wall adjoined each other, or strong intensities appeared on the vaginal area by reflections. Through a semiautomated registration algorithm, images were centered, aligned, and rotated, as shown in the fourth column of Figure 4. Additional segmentation was conducted using overlapped regions of pre– and post–acetic acid application images as a secondary ROI mask. Furthermore, ROI images were sliced based on their clock positions as they were collected for histopathology. Each slice was then used to calculate color distribution histogram, CIE*Lab, and HSV-based Haralick's texture features for VIA.

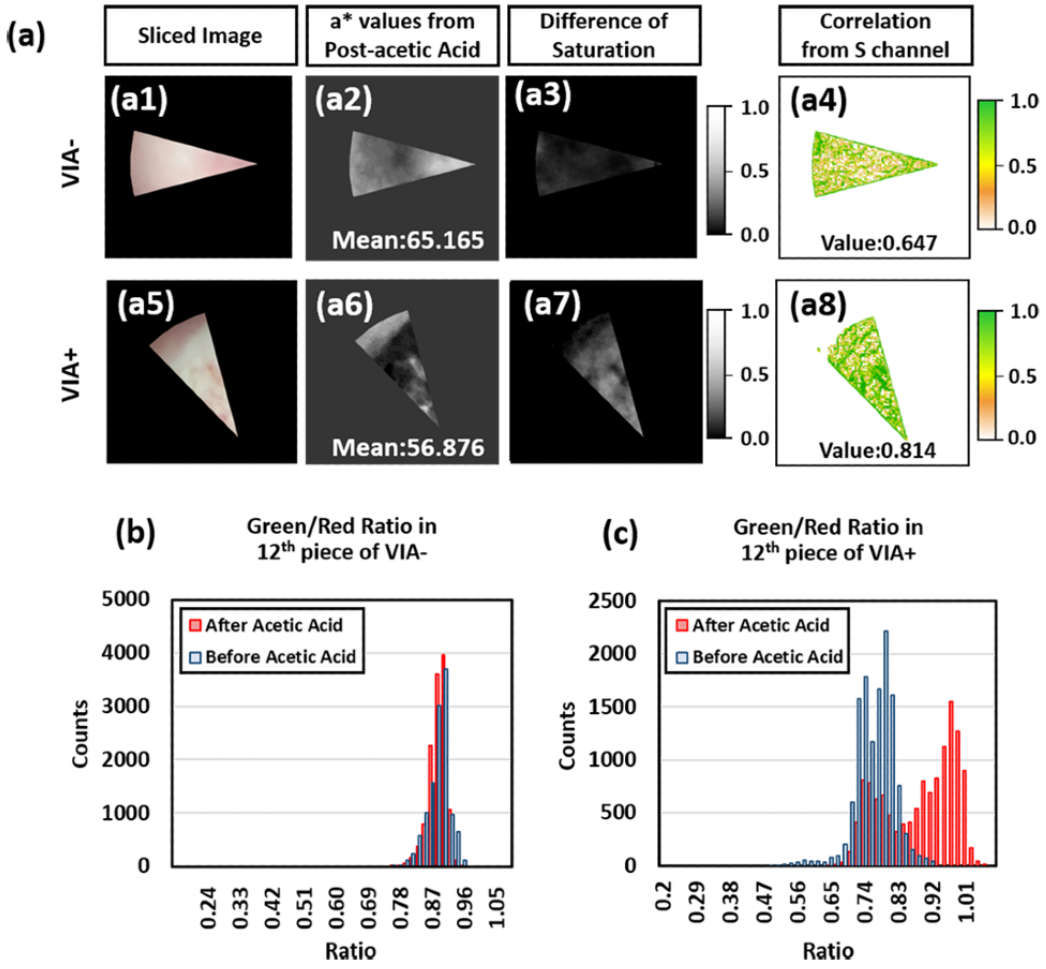
Figure 4. Result of image processing result. SR: speculum reflection; ROI: region of interest.



Feature Extraction

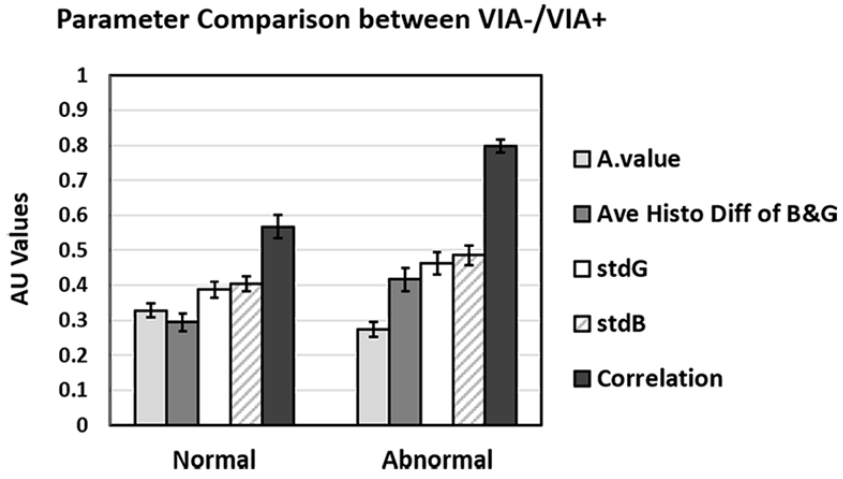
In Figure 5 (a1, a5), representative images of the sliced image are shown for VIA– and VIA+, respectively. Figure 5 (a2, a6) represents a* channel images using post–acetic acid images. Due to aceto-whitening areas, VIA+ showed significantly lower average values in the a* channel from CIE*Lab color space, which represents lower for green and higher for magenta colors. Primarily, lower values are localized in the area where aceto-white exists. In Figure 5 (a3, a7), the images of the difference of saturation channel from pre– to post–acetic acid application are shown. Higher values of saturation were obtained in aceto-whitening regions with VIA+ relative to those with VIA–, where no aceto-whitening regions exist. Using images in Figure 5 (a3, a7), correlation values were computed from Haralick's texture feature and visualized in Figure 5 (a4,8). Contrary to expectation, higher correlation was found in VIA+ compared to VIA– (0-1, usually 1 for higher correlation); however, a significant difference was observed, with a good trend for distinguishing the feature.

Figure 5. Result of feature extraction. (a) Extracted VIA features from sliced VIA-/VIA+ images, showing graphical information in different color spaces. In addition, correlation map was calculated form Haralick's texture feature. (b,c) Representative histogram of green-to-red ratio in VIA-/VIA+. VIA: Visual inspection with acetic acid.



Among the five different features, representative data for the green-to-red ratio were illustrated in Figure 5, part (b) and part (c). There was a relatively small effect from the acetic acid application observed in the green-to-red ratio in the VIA– images, and the histogram of the green-to-red ratio in VIA– shows little change between pre– and post–acetic acid application in terms of distribution and mode. However, the green-to-red ratio increased for VIA+ following application of acetic acid, as shown in the histogram in Figure 5, part (c). Following the application of acetic acid, the distribution of the green intensity histogram broadened along with an increase in intensity. Collecting all the statistical data, we derived the average difference of intensity level in green/blue-to-red ratio and the values of a* channels following application of acetic acid. The standard deviation of green and blue intensity and correlation of Haralick's texture feature were also calculated from the images after the application of acetic acid. Figure 6 summarizes these features with all predictors in abnormal tissues having greater values than those from normal tissues, except for a* values.

Figure 6. Summarization of selected features as predictors for classifying model. VIA: Visual inspection with acetic acid; AU: arbitrary unit.

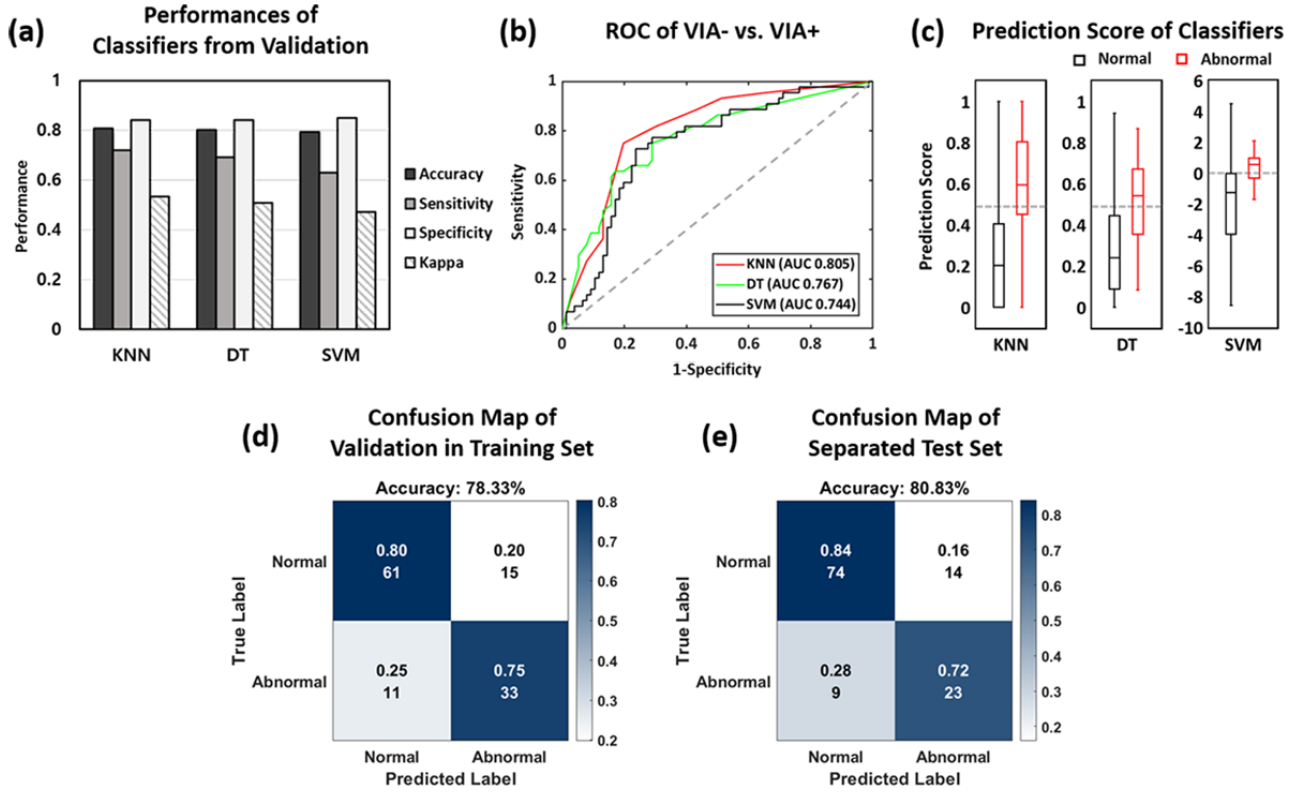


Classification Comparison

By using all the extracted values from 120 sliced image sets, three different types of machine learning classifiers were trained and analyzed. As shown in Figure 7, part (a), KNN yielded the best performance in 10-fold cross-validation. KNN had the best accuracy at 78.3%. with a sensitivity of 75.0%, a specificity of 80.3%, and a Cohen kappa of 0.5423 from 10-fold cross-validation. The second most accurate classifier was DT, with an accuracy of 75.8%, a sensitivity of 63.6%, specificity of 82.9%, and Cohen kappa of 0.4721. SVM's performance was like DT but had the least accuracy for smartphone-based endoscopic VIA, with an accuracy of 74.2%, sensitivity of 72.7%, specificity of 75.0%, and Cohen kappa of 0.4618. In the ROC curve shown in Figure 7, part (b), KNN, DT, and SVM showed an AUC of 0.805, 0.767, and 0.744, respectively. Figure 7, part (c), illustrates the result of prediction scores from each classification on k-fold validation, indicating the probability of an image belonging to either the negative or positive class. From the box plot, separations of the scores of the abnormal and normal data were distinguishable in the KNN model when compared to that of DT and SVM. Most of the abnormal data were distributed over 0.5, but those of normal data were shown under 0.5 in the KNN and DT models. In the SVM model, normal and abnormal data were separated into positive and negative, respectively. As the best performing classifier in our study, the KNN model was selected to evaluate the diagnostic accuracy of VIA.

The confusion matrix for KNN in the validation and test sets is illustrated in Figure 7, parts (d) and (e). The trained KNN model produced similar accuracy for test data to that for validation data, with an accuracy of 80.8%. The sensitivity of 71.9% in the testing data was somewhat lower than that of validation data; however, the specificity of 84.1% was slightly higher than that of the training data.

Figure 7. Comparison & selection of classification model. (a) Performance of 10-fold cross validation in training set comparing the VIA results from each classifier. (b) ROC curves comparing the VIA+ performances. (c) Box plots of prediction scores for each classifying method. (d) Confusion map of validation in training set for KNN model. (e) Confusion map of separated test set for selected KNN model. VIA: visual inspection with acetic acid; ROC: receiver operating characteristic; KNN: k-nearest neighbors.

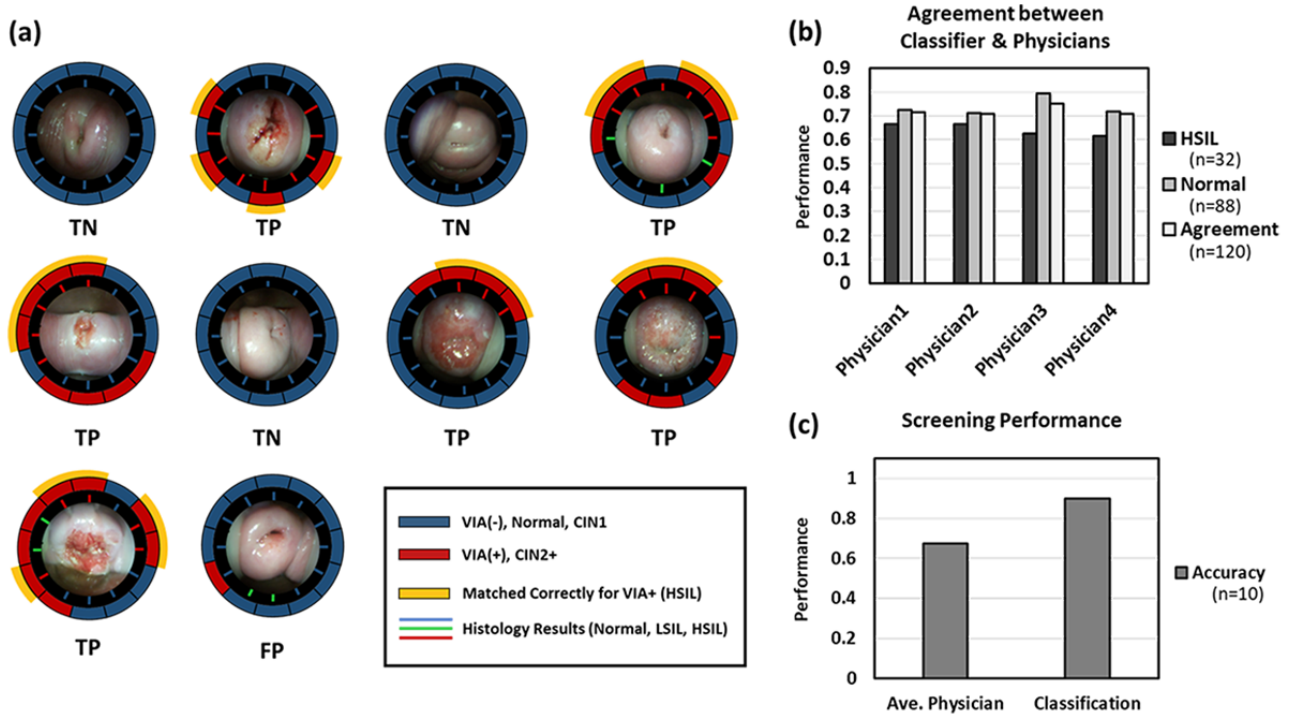


Analysis of the KNN Classification for Smartphone-Based Endoscopic VIA

Figure 8 illustrates the result of the classification with smartphone-based endoscopic VIA images for ten patients that were used for the testing set. By using all the prediction results from the testing data, the locations of abnormal tissue were visualized as a segmented annulus outside of each endoscopic image. The red and blue segments of the annulus indicated VIA+ and VIA– at each clock position, respectively. Out of 120 slice images, 32 were positive and 88 were negative, according to the histopathology. As shown in Figure 8, part (a), all the results of the gold standard for each patient were denoted as colored lines inside the annulus and are aligned along each clock position as a line. Red, green, and blue lines denote abnormal tissues with HSIL (CIN2+), LSIL (CIN1), and normal diagnoses, respectively. When the KNN result was confirmed correctly as VIA+ by the gold standard, segments were labeled yellow on the perimeter of the annulus for each patient. A total of 23 sliced regions were matched correctly for VIA+.

Moreover, a patient was estimated to have an HSIL lesion if there was at least one VIA+ segment included among the 12 locations. Out of a total of 10 patients, six patients were predicted to have HSIL through KNN prediction. Among them, six patients were correctly estimated as VIA+ (true positive). Among the four patients confirmed as VIA–, KNN identified three patients correctly (true negative). However, there were no false negatives predicted by KNN for ten patients. Figure 8, part (b), shows the agreement between the classification algorithm and an individual physician's interpretation for each clock position. The classification algorithm shows a moderate overall agreement ranging from 70.8-75.0%. For binary classifications of patients, the algorithm provides more accurate interpretations, as shown in Figure 8, part (c). The algorithm yielded an accuracy of 90%, and the average accuracy for the physician-achieved accuracy was 68% for ten patients.

Figure 8. Analysis of KNN for smartphone-based endoscopic VIA. (a) Visualization of KNN classifying result for 10 patients with VIA+/VIA-. Lines inside of annulus ring indicate the results of the golden standard for each clock position. Annulus ring presents the result of KNN classification according to each clock position. Outer yellow mark represents the position where both the golden standard and prediction results are matched. (b) Graph presenting the agreement of KNN to each physician. (c) Graph comparing the screening performance for 10 patients interpreted by KNN and physicians. VIA: Visual inspection with acetic acid; KNN: k-nearest neighbors.



Discussion

Principal Results

The findings of this study suggest that the novel smartphone-based approach for VIA with endoscopy and machine learning techniques could potentially be a useful tool for screening cervical cancer in low- and middle-income countries. A smartphone-based endoscope system is a simple and robust screening method that could be performed outside of the office without cost, mobility, and electricity limitations. Here, a moderate overall agreement (Cohen kappa=0.571) was achieved between the interpretations of the smartphone-based endoscopic cervicogram by physicians and the histopathologic results. With the support of machine learning and image processing techniques, the diagnostic performance was evaluated for classifying VIA– or VIA+ patients. We explored three different types of classifiers and selected the KNN algorithm for this study due to its classification performance, and its AUC of 0.805. Both hardware and machine-learned algorithms, based on the gold standard, overperformed compared to the conventional VIA. This was possible because our approach increased the sampling numbers of prediction for each patient compared to the gold standard. From the clinical point of view, finding abnormality in normal tissue is critical, especially in low-resource settings. The locations of abnormal tissue were identified and visualized at each clock position for each patient. With intuitive, perceptual color labels, VIA providers may easily understand where aceto-whitening has appeared and its degree of abnormality. Had our algorithm been used to determine the treatment pathway, the data would effectively assist the VIA providers with decision making.

To the best of our knowledge, this is the first smartphone-based endoscopic VIA work. A smartphone-based endoscopic approach provides some advantages over other smartphone-based VIA studies [14-20,38,39]. Many studies on diagnostic performances of smartphone VIAs have shown promising results, with sufficient accuracy [14-20]. These methods mostly utilized the tripod, or other hardware supports, to obtain pictures, which poses a challenge of limited space for maneuvering during the procedure. Recently, a portable colposcope device with a unique tampon form called the pocket colposcope was introduced with a high concordance to the clinical colposcope. Like our approach, a pocket colposcope enables us to capture images inside of the vagina. However, a pocket colposcope requires a wire-connection to a smartphone as an accessory, with limited hardware variation available [38,39]. Here, using endoscopic probes within a speculum offers noncontact, noninvasive imaging capability, minimizing the cross-contamination risk between patients. Our endoscopic probe enhances the resolution of the images and can achieve more details of mosaicism, punctuation, and aceto-white regions. This could substantially improve the diagnostic accuracy as well as quality control of VIA.

Moreover, various types of endoscopic probes bring flexibility to the system. A smartphone-based endoscope is based on off-the-shelf optics to magnify the images while reducing the signal-to-noise ratio by avoiding high digital zoom. A simple modification of optical components can enable additional functional imaging such as fluorescence, polarization, and multi-spectral imaging, and thus can further improve the visualization of the abnormal tissue [24,40-42]. Furthermore, manipulation of the endoscopic device inside of the vagina is

very intuitive and easy compared to that of conventional approaches that are manipulated outside of the vagina. Thus, neither tripods nor stands are required for acquiring stable images, which in turn enhances the speed of the imaging session.

Image processing algorithms have been developed to minimize external sources of error, such as deviations of illumination power, imaging position, variations of devices, specular reflections, and ROI of the cervix. Our method of ROI segmentation is especially unique, as it uses bit-plane separation and finding overlapped regions to eliminate insufficient information in smartphone-based endoscopic cervicograms. VIA features were mostly extracted based on a statistical calculation of histograms rather than just acquiring the quantitative pixel information from the raw images. This may minimize dependencies on the imaging environment and make image calculation confined to each image, but still afford similar values of computation results when finding features. Extending the RGB images to other color space also provides benefits for calculation. Saturation values from HSV color space afford the extent of purity of the hue data, which results in more intuitive perceptual values between pre– and post–acetic acid application.

Moreover, the saturation value is separated from the brightness value; thus, it is illumination invariant. Empirically, using images of the difference between pre– and post–acetic acid application in the saturation channel for GLCM showed a higher significant difference ($P<.001$) in correlation values between VIA– and VIA+. CIE*Lab color space is also very useful for finding device- or illumination-independent quantitative color measurements. CIE*Lab also separates the luminance component (L*) from the color, which makes color information less sensitive to illumination [28].

Limitations and Future Works

Even though our novel CC screening technique using smartphone-based endoscopic VIA and machine learning demonstrated a positive outcome as a pilot study, much improvement is needed before it can reach its practical potential. First, the accuracy of our algorithm classifying the patients was calculated using a small number of samples. No false-negative data was predicted within a small number of samples. We believe increasing sampling size could additively provide a greater number of sliced data for training and testing classifiers, which could lead to more new cases that eventually lead to much more reliable classification results. Therefore, the overall estimation could be different with a larger data set; however, it will still provide better performance compared to subjective estimation from physicians. Second, our approach for ROI segmentation aimed to precisely distinguish the ectocervix area from the images; however, there are still some portions of vaginal walls left that led to misclassification. Moreover, the location and the area of the cervix are substantially different in each image. For this reason, the image processing algorithm for registering images of pre– and post–acetic acid application was done by manually providing positions of the center and distal ends of the cervical os. Through image standardization by providing the visual guides in software, deviation of positioning and working distance can be minimized and further improve the ROI segmentation performance in an automated processing algorithm. Lastly, our image processing–based classification algorithm was performed on a computer, limiting practical uses in developing countries in its current form. Due to the computational intensity, the current algorithm cannot be operated as a standalone in a smartphone. Recently, Android-based machine learning, as well as image processing techniques, have been introduced in various fields, using SVM, KNN, DT, and Deep Convolutional Neural Networks. As these functions are now available on mobile phones [43-46], we will embed these Android-based machine learning techniques to implement our algorithm on a smartphone alone.

Conclusions

In conclusion, we explored the use of smartphone-based endoscopic VIA and predicted a cervicogram as normal/abnormal using machine learning techniques. Herein, histopathology results, acquired at each clock position on the excised cervical tissue, were provided as ground truth for classifications. VIA features were then extracted after image processing and utilized for training the classifiers. Overall, 120 sliced images obtained from the cervicogram at each clock position were utilized, and three classifiers, such as KNN, DT, and SVM, were compared. Approaches using KNN showed the best performance from holdout 10-fold cross-validation in the training set, with an accuracy of 78.3%, AUC of 0.807, a specificity of 80.3%, and sensitivity of 75.0%. To validate the trained model, we used the other 120 sliced images, achieving an accuracy of 80.8%, specificity of 84.1%, and sensitivity of 71.9%. Prediction values were visualized with intuitive color labels, indicating the normal/abnormal tissue at each clock position for each patient. Calculating the overlapped abnormal tissues between the gold standard and predicted values, our KNN model for classifying VIA–/VIA+ patients overperformed the interpretation results by physicians. Taken together, these results suggest that the smartphone-based endoscopic VIA and analysis based on the machine learning techniques would be a valuable tool for screening VIA in low-resource settings. Moreover, this approach can potentially minimize human subjectivity and be particularly useful in areas where experts or teleconsultations are unavailable.

Acknowledgments

This research was supported by the Samsung Research Funding Center of Samsung Electronics under Project Numbers SRFC-IT1502-54, and the National Research Foundation of Korea grant from the Korean government (No. 2017009566).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Block diagram of image processing and classification. (a) Pre-image processing method, (b) Feature analysis and classification.
[PNG File , 560 KB - mhealth_v8i3e16467_app1.png]

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018 Sep 12;68(6):394-424 [FREE Full text] [doi: 10.3322/caac.21492] [Medline: 30207593]
2. World Health Organization. 2018. Cervical cancer URL: https://www.who.int/cancer/prevention/diagnosis-screening/cervical-cancer/en/ [accessed 2019-01-05]
3. Cohen PA, Jhingran A, Oaknin A, Denny L. Cervical cancer. Lancet 2019 Jan 12;393(10167):169-182. [doi: 10.1016/S0140-6736(18)32470-X] [Medline: 30638582]
4. Huy NVQ, Tam LM, Tram NVQ, Thuan DC, Vinh TQ, Thanh CN, et al. The value of visual inspection with acetic acid and Pap smear in cervical cancer screening program in low resource settings - A population-based study. Gynecol Oncol Rep 2018 May;24:18-20 [FREE Full text] [doi: 10.1016/j.gore.2018.02.004] [Medline: 29527550]
5. Drummond JL, Were MC, Arrossi S, Wools-Kaloustian K. Cervical cancer data and data systems in limited-resource settings: Challenges and opportunities. Int J Gynaecol Obstet 2017 Jul;138 Suppl 1:33-40. [doi: 10.1002/ijgo.12192] [Medline: 28691330]
6. World HO&PAHO. Monitoring national cervical cancer prevention and control programmes: quality control and quality assurance for visual inspection with acetic acid (VIA)-based programmes. Geneva: WHO; 2013.
7. Denny L, Quinn M, Sankaranarayanan R. Chapter 8: Screening for cervical cancer in developing countries. Vaccine 2006 Aug 31;24:71-77. [doi: 10.1016/j.vaccine.2006.05.121] [Medline: 16950020]
8. Raifu AO, El-Zein M, Sangwa-Lugoma G, Ramanakumar A, Walter SD, Franco EL, Congo Screening Study Group. Determinants of Cervical Cancer Screening Accuracy for Visual Inspection with Acetic Acid (VIA) and Lugol's Iodine (VILI) Performed by Nurse and Physician. PLoS One 2017 Jan 20;12(1):e0170631 [FREE Full text] [doi: 10.1371/journal.pone.0170631] [Medline: 28107486]
9. Wright TC. Chapter 10: Cervical cancer screening using visualization techniques. J Natl Cancer Inst Monogr 2003(31):66-71. [doi: 10.1093/oxfordjournals.jncimonographs.a003485] [Medline: 12807948]
10. Bomfim-Hyppólito S, Franco ES, Franco RGDMM, de Albuquerque CM, Nunes GC. Cervicography as an adjunctive test to visual inspection with acetic acid in cervical cancer detection screening. Int J Gynaecol Obstet 2006 Jan;92(1):58-63. [doi: 10.1016/j.ijgo.2005.09.016] [Medline: 16269147]
11. Chen Z, Chen H, Lee T. Use of compact digital cervicography: an adjuvant screening tool for precancerous cervical lesions. Taiwan J Obstet Gynecol 2008 Jun;47(2):187-191 [FREE Full text] [doi: 10.1016/S1028-4559(08)60078-9] [Medline: 18603504]
12. Firnhaber C, Mao L, Levin S, Faesen M, Lewis DA, Goeieman BJ, et al. Evaluation of a cervicography-based program to ensure quality of visual inspection of the cervix in HIV-infected women in Johannesburg, South Africa. J Low Genit Tract Dis 2015 Jan;19(1):7-11 [FREE Full text] [doi: 10.1097/LGT.0000000000000040] [Medline: 24914887]
13. Nam K, Kim S, Sim S, Han S. The performance of tele-cervicography for detection of preinvasive and invasive disease of the uterine cervix as an adjunctive test to Pap smears. Contemp Oncol (Pozn) 2016;20(5):402-406 [FREE Full text] [doi: 10.5114/wo.2016.64604] [Medline: 28373823]
14. Asgary R, Adongo PB, Nwameme A, Cole HVS, Maya E, Liu M, et al. mHealth to Train Community Health Nurses in Visual Inspection With Acetic Acid for Cervical Cancer Screening in Ghana. J Low Genit Tract Dis 2016 Jul;20(3):239-242 [FREE Full text] [doi: 10.1097/LGT.0000000000000207] [Medline: 27030884]
15. 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]
16. Tran PL, Benski C, Viviano M, Petignat P, Combescure C, Jinoro J, et al. Performance of smartphone-based digital images for cervical cancer screening in a low-resource context. Int J Technol Assess Health Care 2018 Jan;34(3):337-342. [doi: 10.1017/S0266462318000260] [Medline: 29921339]
17. Urner E, Delavy M, Catarino R, Viviano M, Meyer-Hamme U, Benski A, et al. A Smartphone-Based Approach for Triage of Human Papillomavirus-Positive Sub-Saharan African Women: A Prospective Study. JMIR Mhealth Uhealth 2017 May 29;5(5):e72 [FREE Full text] [doi: 10.2196/mhealth.6697] [Medline: 28554879]
18. Bagga R, Suri V, Srinivasan R, Khandelwal N, Keswarpu P, Naik S. Feasibility of using mobile smartphone camera as an imaging device for screening of cervical cancer in a low-resource setting. J Postgrad Med Educ Res 2016;50(2):69-74. [doi: 10.5005/jp-journals-10028-1196]

19. Quinley KE, Gormley RH, Ratcliffe SJ, Shih T, Szep Z, Steiner A, et al. Use of mobile telemedicine for cervical cancer screening. J Telemed Telecare 2011;17(4):203-209 [FREE Full text] [doi: 10.1258/jtt.2011.101008] [Medline: 21551217]
20. 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;10(7):e0134309 [FREE Full text] [doi: 10.1371/journal.pone.0134309] [Medline: 26222772]
21. Xu X, Akay A, Wei H, Wang S, Pingguan-Murphy B, Erlandsson B, et al. Advances in Smartphone-Based Point-of-Care Diagnostics. Proc. IEEE 2015 Feb;103(2):236-247. [doi: 10.1109/JPROC.2014.2378776]
22. Liu J, Geng Z, Fan Z, Liu J, Chen H. Point-of-care testing based on smartphone: The current state-of-the-art (2017-2018). Biosens Bioelectron 2019 May 01;132:17-37. [doi: 10.1016/j.bios.2019.01.068] [Medline: 30851493]
23. Im H, Castro CM, Shao H, Liong M, Song J, Pathania D, et al. Digital diffraction analysis enables low-cost molecular diagnostics on a smartphone. Proc Natl Acad Sci U S A 2015 May 05;112(18):5613-5618 [FREE Full text] [doi: 10.1073/pnas.1501815112] [Medline: 25870273]
24. Grant BD, Quang T, Possati-Resende JC, Scapulatempo-Neto C, de Macedo Matsushita G, Mauad EC, et al. A mobile-phone based high-resolution microendoscope to image cervical precancer. PLoS One 2019;14(2):e0211045 [FREE Full text] [doi: 10.1371/journal.pone.0211045] [Medline: 30726252]
25. Bae JK, Vavilin A, You JS, Kim H, Ryu SY, Jang JH, et al. Smartphone-Based Endoscope System for Advanced Point-of-Care Diagnostics: Feasibility Study. JMIR Mhealth Uhealth 2017 Jul 27;5(7):e99 [FREE Full text] [doi: 10.2196/mhealth.7232] [Medline: 28751302]
26. Park SY, Follen M, Milbourne A, Rhodes H, Malpica A, MacKinnon N, et al. Automated image analysis of digital colposcopy for the detection of cervical neoplasia. J Biomed Opt 2008;13(1):014029 [FREE Full text] [doi: 10.1117/1.2830654] [Medline: 18315387]
27. Ramapraba PS, Chitra MP, Prem Kumar M. Effective lesion detection of colposcopic images using active contour method. Biomed Res 2017 Feb 23;28:S255-S264 [FREE Full text]
28. Asiedu MN, Simhal A, Chaudhary U, Mueller JL, Lam CT, Schmitt JW, et al. Development of Algorithms for Automated Detection of Cervical Pre-Cancers With a Low-Cost, Point-of-Care, Pocket Colposcope. IEEE Trans Biomed Eng 2019 Aug;66(8):2306-2318 [FREE Full text] [doi: 10.1109/TBME.2018.2887208] [Medline: 30575526]
29. Li W, Van Raad V, Gu J, Hansson U, Hakansson J, Lange H. Computer-Aided Diagnosis (CAD) for cervical cancer screening and diagnosis: A new system design in medical image processing. : Springer-Verlag; 2005 Presented at: Proceedings of the First international conference on Computer Vision for Biomedical Image Applications; 21 Oct; Beijing, China p. 240-250. [doi: 10.1007/11569541_25]
30. Liu J, Peng Y, Zhang Y. A Fuzzy Reasoning Model for Cervical Intraepithelial Neoplasia Classification Using Temporal Grayscale Change and Textures of Cervical Images During Acetic Acid Tests. IEEE Access 2019;7:13536-13545. [doi: 10.1109/access.2019.2893357]
31. Fernandes K, Cardoso JS, Fernandes J. Automated Methods for the Decision Support of Cervical Cancer Screening Using Digital Colposcopies. IEEE Access 2018;6:33910-33927. [doi: 10.1109/access.2018.2839338]
32. Ohri N, Kapoor C, Mohemmed RP, Vaidya S. An update on intraoral application of colposcopy. J Oral Maxillofac Pathol 2014;18(3):403-410 [FREE Full text] [doi: 10.4103/0973-029X.151328] [Medline: 25948996]
33. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Trans. Syst., Man, Cybern 1973 Nov;SMC-3(6):610-621. [doi: 10.1109/tsmc.1973.4309314]
34. Sankaranarayanan R, Wesley R. A Practical Manual on Visual Screening for Cervical Neoplasia. IARC Technical Publication, No. 41. Lyon, France: IARC Press; 2003.
35. Selvarajah S, Kodituwakku SR. Analysis and comparison of texture features for content based image retrieval. Int J Latest Trends Comput 2011 Mar;2(1):108-113 [FREE Full text]
36. Albregtsen F. Statistical Texture Measures Computed from Gray Level Cooccurrence Matrices. In: Image Processing Laboratory. University of Oslo: Department of Informatics; 1995:1-14.
37. Refaeilzadeh P, Tang L, Liu H. Cross-validation. In: Encyclopedia of Database Systems. Boston, MA: Springer; 2009:532-538.
38. Lam CT, Mueller J, Asma B, Asiedu M, Krieger MS, Chitalia R, et al. An integrated strategy for improving contrast, durability, and portability of a Pocket Colposcope for cervical cancer screening and diagnosis. PLoS One 2018;13(2):e0192530 [FREE Full text] [doi: 10.1371/journal.pone.0192530] [Medline: 29425225]
39. Mueller JL, Asma E, Lam CT, Krieger MS, Gallagher JE, Erkanli A, et al. International Image Concordance Study to Compare a Point-of-Care Tampon Colposcope With a Standard-of-Care Colposcope. J Low Genit Tract Dis 2017 Apr;21(2):112-119 [FREE Full text] [doi: 10.1097/LGT.0000000000000306] [Medline: 28263237]
40. Ren W, Qu Y, Pei J, Xiao L, Zhang S, Chang S. Development of a multimodal colposcopy for characterization of cervical intraepithelial neoplasia. ASME J Med Devices 2017 Sep;11(3):031005. [doi: 10.1115/1.4036335]
41. Meena BL, Singh P, Sah AN, Pandey K, Agarwal A, Pantola C, et al. Intrinsic fluorescence for cervical precancer detection using polarized light based in-house fabricated portable device. J Biomed Opt 2018 Jan;23(1):1-7 [FREE Full text] [doi: 10.1117/1.JBO.23.1.015005] [Medline: 29341542]

42. Vizet J, Rehbinder J, Deby S, Roussel S, Nazac A, Soufan R, et al. In vivo imaging of uterine cervix with a Mueller polarimetric colposcope. Sci Rep 2017 Jun 01;7(1):2471-2412 [FREE Full text] [doi: 10.1038/s41598-017-02645-9] [Medline: 28572602]
43. Aleem M, Hameed N, Anjum A. m-Skin Doctor: A Mobile Enabled System for Early Melanoma Skin Cancer Detection Using Support Vector Machine. In: eHealth 360°: International Summit on eHealth, Budapest, Hungary, 2016 June 14-16, Revised Selected Papers. Cham, Switzerland: Springer; 2017:468-475.
44. Kim H, Awofeso O, Choi S, Jung Y, Bae E. Colorimetric analysis of saliva–alcohol test strips by smartphone-based instruments using machine-learning algorithms. Appl. Opt 2016 Dec 23;56(1):84-92. [doi: 10.1364/ao.56.000084]
45. Ramlakhan K, Shang Y. A mobile automated skin lesion classification system. In: Proceedings of the 23rd International Conference on Tools with Artificial Intelligence. Washington: IEEE; 2011 Presented at: 2011 IEEE; 2011 7-9 Nov; Boca Raton, FL, USA p. 138-141. [doi: 10.1109/ICTAI.2011.29]
46. Ech-Cherif A, Misbhauddin M, Ech-Cherif M. Deep Neural Network based mobile dermoscopy application for triaging skin cancer detection. In: 2nd International Conference on Computer Applications & Information Security.: IEEE; 2019 Presented at: ICCAIS; 2019 May 1-3; Riyadh, Saudi Arabia p. 1-6. [doi: 10.1109/CAIS.2019.8769517]

Abbreviations

3D: three-dimensional
AU: arbitrary units
AUC: area under the curve
CC: cervical cancer
CIE: Commission Internationale de l'Eclairage
CIN: cervical intraepithelial neoplasia
CV: cross-validation
DT: Decision tree
GLCM: gray-level cooccurrence matrix
HSIL: high-grade squamous intraepithelial lesion
HSV: Hue, Saturation, Value
KNN: k-nearest neighbors
LEEP: loop electrosurgical excision procedure
LSIL: low-grade squamous intraepithelial lesion
RGB: red, green, blue
ROC: receiver operating characteristic
ROI: region of interest
SVM: support vector machine
USAF: Unites States air force
VIA: visual inspection with acetic acid

Edited by G Eysenbach; submitted 24.10.19; peer-reviewed by JY Hwang, D Carvalho; comments to author 19.11.19; revised version received 10.01.20; accepted 27.01.20; published 11.03.20.

Please cite as:
Bae JK, Roh HJ, You JS, Kim K, Ahn Y, Askaruly S, Park K, Yang H, Jang GJ, Moon KH, Jung W
Quantitative Screening of Cervical Cancers for Low-Resource Settings: Pilot Study of Smartphone-Based Endoscopic Visual Inspection After Acetic Acid Using Machine Learning Techniques
JMIR Mhealth Uhealth 2020;8(3):e16467
URL: http://mhealth.jmir.org/2020/3/e16467/
doi:10.2196/16467
PMID:32159521

Original Paper

Effectiveness of a Mobile eHealth App in Guiding Patients in Pain Control and Opiate Use After Total Knee Replacement: Randomized Controlled Trial

Yvette Pronk[1], MSc; Maud Cornelia Wilhelmina Maria Peters[1], MSc; Amarsing Sheombar[2], MD; Justus-Martijn Brinkman[3], MD, PhD

[1]Research Department, Kliniek ViaSana, Mill, Netherlands

[2]Department of Anaesthesiology, Kliniek ViaSana, Mill, Netherlands

[3]Department of Orthopaedic Surgery, Kliniek ViaSana, Mill, Netherlands

Corresponding Author:
Yvette Pronk, MSc
Research Department
Kliniek ViaSana
Hoogveldseweg 1
Mill, 5451 AA
Netherlands
Phone: 31 485476330
Email: y.pronk@viasana.nl

Abstract

Background: Little is known about pain and opiate use at home directly after total knee replacement (TKR). Due to adverse effects, low opiate use is desired. An electronic health app (PainCoach) was developed to guide patients in pain control and opiate use.

Objective: The aim of this paper was to investigate the effects of the PainCoach app on pain control and opiate use in patients who underwent TKR during the first 2 weeks at home after surgery.

Methods: In an unblinded randomized controlled trial, patients scheduled for TKR were offline recruited and randomized to a PainCoach group or control group. In the PainCoach group, the PainCoach app was downloaded on each patient's smartphone or tablet. In response to the patient's input of the pain experienced, the PainCoach app gave advice on pain medication use, exercises/rest, and when to call the clinic. This advice was the same as that received during usual care. The control group received usual care. The primary outcomes were opiate use and visual analog scale (VAS) pain scores at rest, during activity, and at night during the first 2 weeks at home after surgery, which were collected daily from day 1 until 14 postoperatively by online questionnaires. The actual amount of app use was recorded, and active use was defined as ≥12 total app uses.

Results: The pain scores did not differ between the groups. The PainCoach group (n=38) used 23.2% less opiates (95% CI −38.3 to −4.4; *P*=.02) and 14.6% more acetaminophen (95% CI 8.2-21.3; *P*<.001) when compared with the findings in the control group (n=33). The PainCoach app was used 12 (IQR 4.5-22.0) times per patient. In the active PainCoach subgroup (n=19), the following were noted when compared with the findings in the control group: 4.1 times faster reduction of the VAS pain score during activity (95% CI −7.5 to −0.8; *P*=.02), 6.3 times faster reduction of the VAS pain score at night (95% CI −10.1 to −2.6; *P*=.001), 44.3% less opiate use (95% CI −59.4 to −23.5; *P*<.001), 76.3% less gabapentin use (95% CI −86.0 to −59.8; *P*<.001), and 21.0% more acetaminophen use (95% CI 12.6-30.0; *P*<.001).

Conclusions: The use of the PainCoach app contributes to reduced opiate use in the initial period at home after TKR. Active use of this app leads to a further reduction in opiate use and improved pain control.

Trial Registration: ClinicalTrials.gov NCT03961152; https://clinicaltrials.gov/ct2/show/NCT03961152

(JMIR Mhealth Uhealth 2020;8(3):e16415) doi:10.2196/16415

KEYWORDS

opiate use; pain; total knee replacement; app; eHealth

Introduction

Total knee replacement (TKR) is a successful treatment option for patients with end-stage knee osteoarthritis (OA) [1]. Moderate-to-severe pain after TKR can be expected [2,3]. Local infiltration anesthesia (LIA) techniques and so-called fast-track recovery programs have resulted in reduced pain and early mobilization, subsequently reducing the length of stay in hospital and increasing patient satisfaction [4-7]. Previous research established several factors associated with increased pain after TKR [8-19] (Table 1). Postoperative pain inhibits recovery, increases morbidity, and may result in chronic pain, ultimately limiting the effectiveness of TKR [6,20]. Therefore, pain should be controlled optimally both in the hospital and at home.

Table 1. Factors associated with increased pain after total knee replacement.

Factor	Association with increased pain after TKR[a]
Gender	Being female [8-12]
Age	Older age [8,10,13]
BMI[b]	Higher BMI[b] [8,10]
ASA[c] score	Higher ASA[c] score [10]
Pain catastrophization	Higher pain catastrophization score [12,14-17]
Comorbidity	Presence of comorbidities [8,10,13,18]
Previous knee surgery	Having a history of knee surgery [10]
Preoperative pain	Higher preoperative pain severity [8,12,18,19]
Social support	Poor social support [13]
Preoperative mental health	Poor preoperative mental health [8,10,13,18]

[a]TKR: total knee replacement.

[b]BMI: body mass index.

[c]ASA: American Society of Anesthesiologists.

Although pain is usually under control during hospital stay, less is known about pain control in the initial period at home after TKR. Current pain management strategies include a combination of nonsteroidal anti-inflammatory drugs (NSAIDs), nonnarcotic medication, opiates, and exercise [4]. Although opiates are very effective for reducing pain, serious adverse effects, such as nausea, itching, reduced gut mobility, and urinary retention, often occur [21]. Addiction to opiates is an ever increasing problem and may ultimately lead to an increased risk of death [22]. The amount of opiate use should therefore be kept to a minimum. Orthopedic surgery, however, accounts for an estimated 8.8% of prolonged prescription opiate use [23]. Therefore, alternative pain management strategies are needed. Electronic health (eHealth) apps can be used to guide patients in improving their pain management strategies at home. An important benefit of these apps is that patients can access the information provided directly and anywhere whenever necessary [24-29]. The number of older adults with internet access and acceptance of internet-based interventions is increasing, and patients tend to remember up to 80% of the information acquired from interactive education [30,31].

With this in mind, to manage pain better and potentially decrease opiate use, an eHealth app named PainCoach was developed. This app aims to help patients control their pain better in the initial period at home after TKR, including optimal use of the available pain medication. This study aimed to determine the effects of PainCoach on pain control and opiate use in TKR patients in the first 2 weeks at home after surgery. The hypothesis was that the use of this app would decrease pain and opiate use.

Methods

Study Design

An unblinded, randomized, controlled, single-center trial was performed at Kliniek ViaSana (Mill, The Netherlands). Patients with an American Society of Anesthesiologists (ASA) score of I-II, a body mass index (BMI) of ≤35, and a plan to undergo primary TKR between February and June 2016 were enrolled. Four experienced high-volume knee surgeons performed all surgeries, and three experienced anesthesiologists administered spinal anesthesia. The same type of TKR implant was used in all patients (NexGen LPS, ZimmerBiomet, Warsaw, Indiana). All surgeries were performed using a tourniquet. The pain management protocol consisted of preoperatively administered medication, LIA injections during surgery directly before cementing the implant, and a step-wise postoperative pain management protocol (Multimedia Appendix 1). Patients were excluded if they did not possess a smartphone or tablet, had a contraindication to any of the medications used in the study, did not have an email address, did not have internet at home, did not have a thorough command of the Dutch language, had memory disorders, or had surgery under general anesthesia. Patients were recruited over the phone by the research staff after being scheduled for primary TKR under spinal anesthesia, and contraindication to any of the medications used in the study and presence of memory disorders were checked by the anesthesiologists. Patients were asked over the phone if they possessed a smartphone or tablet, had an email address, had internet at home, and had a thorough command of the Dutch language. Patient information and informed consent were sent

by postal service if a patient met the criteria and was interested to participate. Patients were considered lost to follow-up if they completed less than two postoperative questionnaires during the first 2 weeks at home. Power analysis (significance level: .05, power: 90%) showed that 35 patients would be needed in each group to detect a difference of 10 points on a visual analog scale (VAS) for pain (VAS pain, 0-100). Written informed consent was obtained from all participants. The study was approved by the medical ethics committee of St. Anna Hospital (Geldrop, The Netherlands, Study ID: 5.12) and was registered at Clinicaltrials.gov retrospectively (ID: NCT03961152).

Randomization

Included unblinded patients were randomly assigned to the PainCoach or control group using lots presented in sealed opaque envelopes during admission. All lots were created and sealed by a researcher in the ratio of 1:1. A blinded nurse presented the envelops to a patient, and the patient selected one to complete randomization. All patients received the usual pain management care including pre-, peri-, and postoperative pain medication (Multimedia Appendix 1), participated in group information meetings, received an information booklet, and could contact the clinic at any time (24 hours a day/7 days a week) in case of any remaining questions. In the PainCoach group, in addition to receiving the aforementioned usual care, the PainCoach app (Interactive Studios, Rosmalen, The Netherlands) was downloaded on each patient's smartphone or tablet, using a unique download code. In this way, the PainCoach app was not available to the control group. An unblinded nurse provided the code and assisted the patient by completing the download process of the app during admission. The app gave the same advice as that during usual care. After only entering the date of surgery as patient data, the app allowed patients to input their pain level (no pain, bearable pain, unbearable pain, or untenable pain) whenever they wanted until day 14 after surgery. Based on the patient's input and taking into account the number of days after surgery, the app provided advice on pain medication use, physiotherapy exercises including videos, use of ice or heat packs, rest, immobilization of the operated leg, and when to call the clinic (Multimedia Appendix 2). Patients in the PainCoach group were not subjected to any treatment that was different from that in the control group (ie, advice on pain management was delivered in an extra and different way, but the pain medication itself was exactly the same for both groups). During the study, no major changes or revisions were made to the PainCoach app.

Outcomes and Measurements

Beside the actual amount of app use, all the outcome measurements were assessed using a digital, online, automated collection system (OnlinePROMs, Interactive Studios, Rosmalen, The Netherlands), which automatically sent an invitation by email to complete an online questionnaire preoperatively, daily from day 1 to 14, and at 1 month postoperatively. In case of nonresponse to the preoperative or 1-month questionnaire, an automatic reminder was sent after 3 days. The invitation to complete the daily questionnaire was sent at 5 pm, and patients had access to the questionnaire until midnight.

The primary outcomes were opiate use and pain score of the operated knee at rest, during activity, and at night in the first 2 weeks at home after TKR. The pain score was measured on a VAS for pain, which ranged from 0 (no pain) to 100 (worst imaginable pain), preoperatively, daily from day 1 to 14, and at 1 month postoperatively [32-35]. Severe pain was defined as a VAS pain score from 70 to 100. Opiate (oxycodon; 5 mg per tablet; different manufacturers) use was recorded in quantities per 24 hours from day 1 to 14.

The secondary outcomes in the first 2 weeks at home and 1 month after TKR included other pain medication use (ie, NSAIDs [diclofenac], acetaminophen, or gabapentin; different manufacturers), which was also recorded in quantities per 24 hours from day 1 to 14. Additionally, pain acceptance at rest, during activity, and at night was assessed with a happy smiley (acceptable pain) and a sad smiley (unacceptable pain) preoperatively, daily from day 1 to 14, and at 1 month postoperatively. Experiences with the executed recommended physiotherapy exercises were recorded daily from day 1 to 14 on a 3-item scale (did too much, exactly enough, or could have done more exercises). Moreover, function and quality of life were measured preoperatively and 1 month postoperatively. Knee function was assessed using the Knee Injury and Osteoarthritis Outcome Score–Physical Function Short-form (KOOS-PS) on a scale from 0 (no difficulty) to 100 (extreme difficulty) [36]. The Oxford Knee Score was used to measure combined function and pain on a scale from 0 (most severe symptoms) to 48 (least severe symptoms) [37]. Quality of life was measured using the EuroQol-5 Dimensions (EQ-5D) 3-level version (EQ-5D-3L) questionnaire consisting of the following two scores: EQ VAS score, which is assessed on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state), and EQ-5D descriptive system [38]. The PainCoach app's perceived effectiveness (usability, added value, and likelihood of being recommended to others) was recorded on a 5-item scale ranging from totally agree to totally disagree at day 14 after surgery. Each downloaded app had its own app code that was used to record the actual amount of app use. As the admission period was generally 1 or 2 days, outcomes were measured until day 14 after surgery, and outcomes at home were investigated, the outcome active PainCoach app use was defined as using the app at least 12 times in total.

Preoperative opiate and other pain medication use, age, gender, ASA score, BMI, preoperative comorbidities, history of knee surgery on the same side, Charnley score, date of surgery, date of discharge, and complication data were collected from the electronic patient records. Pain coping, anxiety, education level, and marital status were determined preoperatively using an online questionnaire. Pain coping was measured using the pain coping and cognition list scored from 1 (totally disagree) to 6 (totally agree), and it had the following four categories: catastrophizing, pain coping, internal pain management, and external pain management [39].

Statistical Analysis

Analysis was performed using SPSS version 25.0 (IBM Corp, Armonk, New York). All measured outcomes from day 1 until day 14 after surgery were recoded into measured outcomes for

days at home by subtraction of the admission period. Patient characteristics were analyzed using descriptive statistics, and data were checked for normal distribution. Differences in mean, median, or percentage were tested using the independent two-sample *t*-test, Mann-Whitney *U* test, likelihood analysis, Fisher's test, or Pearson's chi-squared test, depending on the type of data. Mixed linear models were used to analyze the overall rate of decrease or increase for continuous data, and generalized linear models were used to analyze the percentage decrease or increase for count and nominal data. Additional analysis was performed to compare the active PainCoach subgroup with the control group, with correction for differences in preoperative data. Statistical significance was set at $P<.05$, and trends were defined as $.05<P<.10$.

Results

Patient Characteristics

A total of 97 patients were included, and of these, 76 patients were randomized. Because of loss to follow-up, the final analysis was performed with 71 patients (PainCoach group, n=38; control group, n=33) (Figure 1). The response rates for the daily questionnaires at home were 91% in the PainCoach group and 89% in the control group.

No statistically significant differences in patient characteristics were found between the PainCoach group and control group. The preoperative VAS pain score at night was significantly lower in the active PainCoach subgroup (n=19) than in the control group (P=.02) (Table 2).

Figure 1. Study flowchart. IOS: iPhone operating system; TKR: total knee replacement.

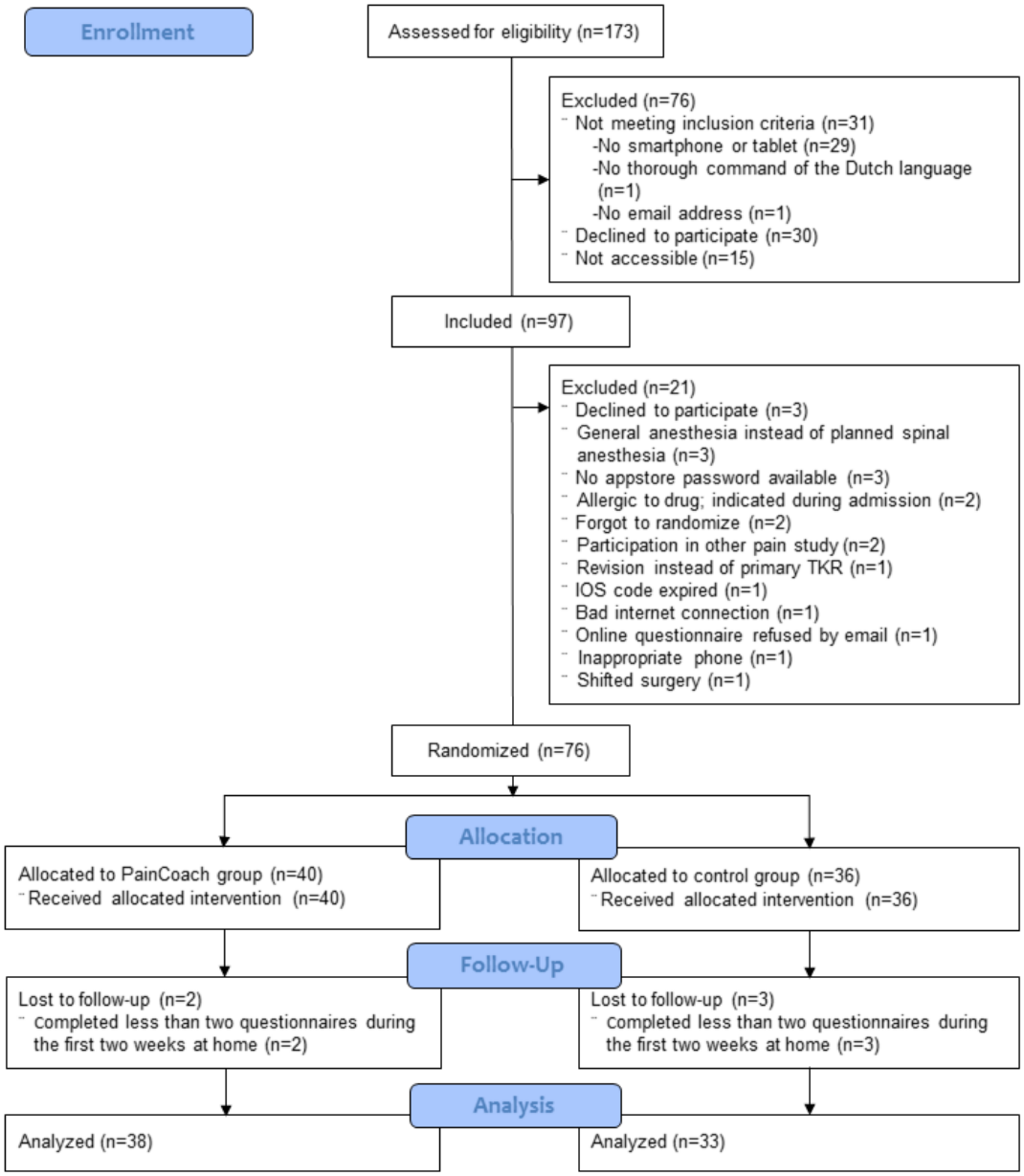


Table 2. Characteristics of patients in the PainCoach group, active PainCoach subgroup, and control group.

Characteristic	1. PainCoach (n=38)	2. Active PainCoach (n=19)	3. Control (n=33)	*P* value (1 vs 3)	*P* value (2 vs 3)
Gender (male), n (%)	23 (61)	13 (68)	19 (58)	.80	.44
Age (years), mean (SD)	62.6 (7.0)	62.8 (6.1)	64.6 (7.5)	.24	.38
BMI[a], mean (SD)	27.6 (3.5)	26.7 (3.4)	27.8 (3.0)	.83	.24
ASA[b] (I), n (%)	18 (47)	11 (58)	12 (36)	.35	.13
Preoperative comorbidities, n (%)	14 (37)	8 (42)	17 (52)	.21	.51
Preoperative prescription, n (%)					
NSAIDs[c]	5 (13)	3 (16)	6 (18)	.20	.48
Acetaminophen	1 (3)	0 (0)	1 (3)	.92	>.99
Opiate	3 (8)	0 (0)	0 (0)	.24	>.99
Gabapentin	0 (0)	0 (0)	0 (0)	>.99	>.99
Preoperative anxiety, n (%)				.59	>.99
No anxiety	33 (87)	18 (95)	30 (91)		
Some anxiety	5 (13)	1 (5)	3 (9)		
Much anxiety	0 (0)	0 (0)	0 (0)		
History of knee surgery on the same side, n (%)	27 (71)	15 (79)	21 (64)	.51	.25
Charnley score, n (%)				.64	.98
One knee affected with OA[d]	22 (58)	12 (63)	19 (58)		
Both knees affected with OA	7 (18)	3 (16)	6 (18)		
Contralateral TKR[e]	5 (13)	1 (5)	2 (6)		
Multiple joints affected with OA	4 (11)	3 (16)	6 (18)		
Education level, n (%)				.33	.30
Primary school	3 (8)	1 (5)	1 (3)		
Secondary school	14 (37)	7 (37)	10 (31)		
Tertiary school	21 (55)	11 (58)	21 (66)		
Marital status, n (%)				.19	.09
Married	29 (76)	18 (95)	22 (67)		
Other[f]	9 (24)	1 (5)	11 (33)		
Pain coping, mean (SD)					
Catastrophization	2.5 (0.7)	2.5 (0.8)	2.3 (0.6)	.15	.32
Pain coping	3.6 (1.0)	3.8 (1.0)	3.7 (0.8)	.68	.66
Internal pain management	4.1 (0.8)	4.2 (0.9)	3.9 (0.8)	.24	.29
External pain management	2.7 (0.8)	2.7 (0.7)	2.5 (0.8)	.31	.36
Preoperative VAS[g] pain, median (IQR[h])					
Knee at rest	33.0 (20.8-52.8)	33.0 (13.0-43.0)	32.0 (17.8-49.0)	.65	.82
Knee during activity	60.5 (36.5-77.3)	57.0 (30.0-75.0)	60.0 (43.3-73.8)	.82	.69
Knee at night	20.5 (4.8-42.5)	15.0 (1.0-30.0)	35.5 (15.0-58.5)	.11	.02[i]
Preoperative acceptable pain, n (%)					
Knee at rest	29 (76)	16 (84)	27 (82)	.40	>.99
Knee during activity	16 (42)	10 (53)	13 (39)	.90	.41
Knee at night	28 (74)	16 (84)	27 (82)	.28	>.99

Characteristic	1. PainCoach (n=38)	2. Active PainCoach (n=19)	3. Control (n=33)	*P* value (1 vs 3)	*P* value (2 vs 3)
Preoperative KOOS-PS[j], median (IQR)	47.3 (41.6-55.3)	46.1 (40.3-54.4)	48.5 (40.3-57.9)	.76	.96
Preoperative OKS[k], mean (SD)	25.3 (7.2)	27.0 (7.2)	24.8 (5.6)	.75	.23
Preoperative EQ-5D[l] descriptive system, median (IQR)	0.775 (0.471-0.783)	0.775 (0.516-0.807)	0.775 (0.651-0.807)	.27	.81
Preoperative EQ VAS[m], median (IQR)	86.0 (73.6-94.3)	87.0 (79.0-93.0)	86.0 (74.0-95.5)	.89	.72
Complications, n (%)	3 (8)	2 (11)	1 (3)	.62	.55

[a]BMI: body mass index.

[b]ASA: American Society of Anesthesiologists.

[c]NSAIDs: nonsteroidal anti-inflammatory drugs.

[d]OA: osteoarthritis.

[e]TKR: total knee replacement.

[f]Other marital status: single, living together, divorced, widow(er), living apart together relationship, different.

[g]VAS: visual analog scale.

[h]IQR: interquartile range.

[i]Significant difference ($P<.05$).

[j]KOOS-PS: Knee Injury and Osteoarthritis Outcome Score–Physical Function Short-form.

[k]OKS: Oxford Knee Score.

[l]EQ-5D: EuroQol-5 Dimensions.

[m]EQ VAS: EuroQol visual analog scale.

Visual Analog Scale Pain Scores and Opiate Use

During the first 2 weeks at home, the PainCoach group had VAS pain scores of 17.0 (IQR 5.0-30.0) at rest, 20.0 (IQR 7.0-35.0) during activity, and 17.0 (IQR 4.0-37.0) at night. The control group had VAS pain scores of 20.0 (IQR 7.0-33.0) at rest, 21.0 (IQR 10.0-38.0) during activity, and 20.5 (IQR 8.0-40.0) at night. Pain was classified as severe on one or more days in 21% (8/38) of patients from the PainCoach group and 30% (10/33) of patients from the control group. No statistically significant differences were found between the two groups in terms of the VAS pain scores at rest, during activity, and at night (Figure 2A-C, Table 3). Regarding opiate use, the PainCoach group used a mean of 0.4 (SD 0.7) tablets a day and the control group used a mean of 0.5 (SD 0.8) tablets a day. Opiate use was significantly reduced by 23.2% in the PainCoach group when compared with the finding in the control group (95% CI −38.3 to −4.4; $P=.02$) (Figure 2A-C, Table 3). One month after surgery, no statistically significant differences in the VAS pain scores were found between the PainCoach group and control group (Table 4).

Adjusted analyses showed that the active PainCoach subgroup had VAS pain scores of 10.0 (IQR 4.0-26.3) at rest, 12.0 (IQR 5.0-25.0) during activity, and 10.0 (IQR 2.8-28.0) at night during the first 2 weeks at home. Pain was reported as severe on one or more days in 16% (3/19) of patients from the active PainCoach subgroup. The VAS pain score during activity significantly decreased 4.1 times faster in the active PainCoach subgroup when compared with the finding in the control group (95% CI −7.5 to −0.8; $P=.02$) (Figure 2E, Table 3). The VAS pain score at night significantly decreased 6.3 times faster in the active PainCoach subgroup when compared with the finding in the control group (95% CI −10.1 to −2.6; $P=.001$) (Figure 2F, Table 3). The mean opiate use was 0.3 (SD 0.5) tablets a day in the active PainCoach subgroup. Opiate use was significantly reduced by 44.3% in the active PainCoach subgroup when compared with the finding in the control group (95% CI −59.4 to −23.5; $P<.001$) (Figure 2D-F, Table 3). One month after surgery, no statistically significant differences in VAS pain scores were found between the active PainCoach subgroup and control group (Table 4).

Figure 2. VAS pain scores and opiate use in the PainCoach group and control group at rest (A), during activity (B), and at night (C) and in the active PainCoach subgroup and control group at rest (D), during activity (E), and at night (F) on separate days and in the overall first period at home. a: significant difference in VAS pain (P<.05); b: significant difference in opiate use (P<.05); c: trend in VAS pain (.05<P<.10); VAS: visual analog scale.

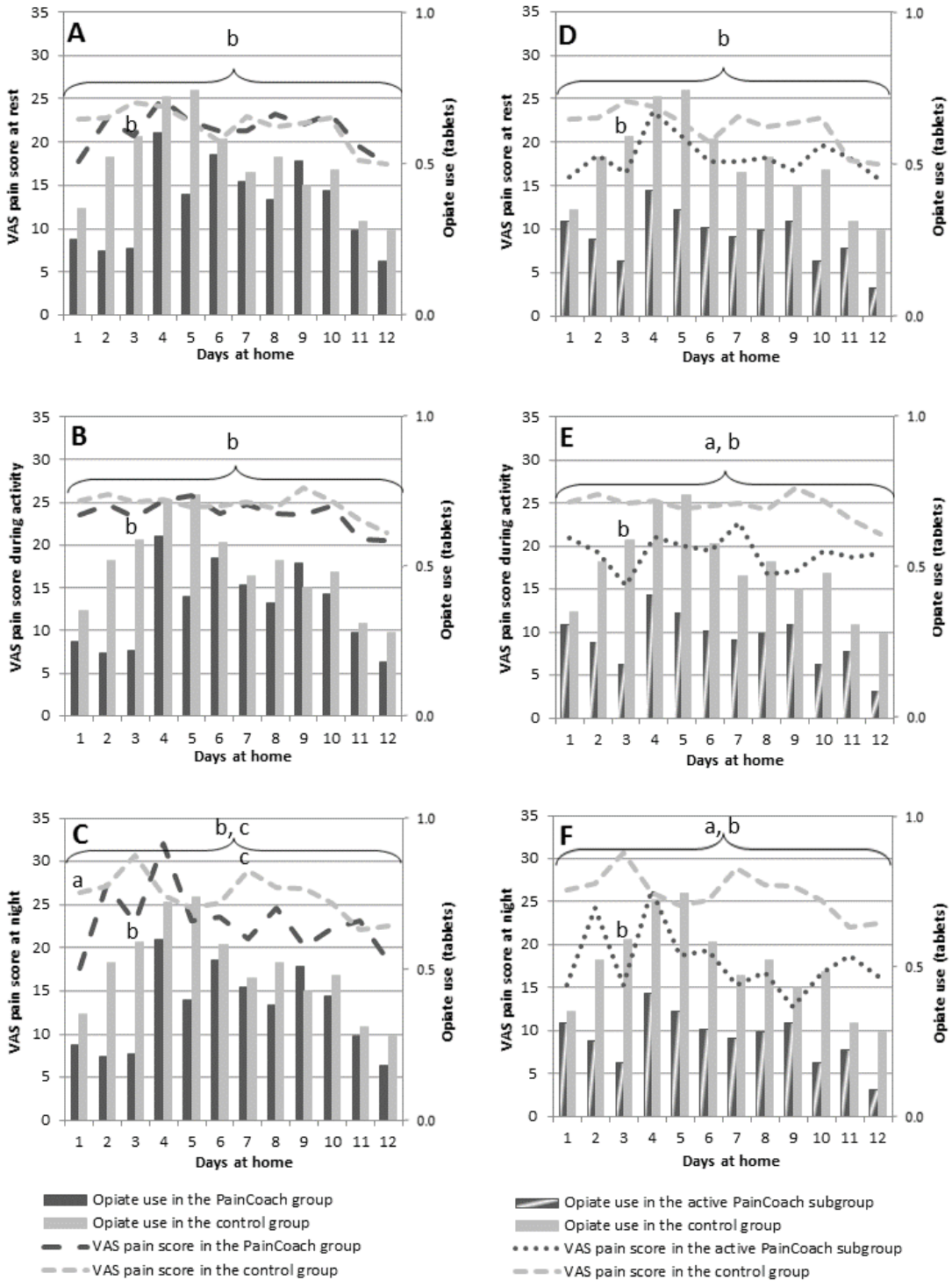


Table 3. Findings in the PainCoach group, active PainCoach subgroup, and control group during the first 2 weeks at home.

Variable	1. PainCoach versus control	2. Active PainCoach versus control	*P* value (1)	*P* value (2)
VAS[a] pain, decrease or increase (rate)				
Knee at rest	↓0.3	↓1.9	.86	.27
Knee during activity	↓1.0	↓4.1	.48	.02[b]
Knee at night	↓3.0	↓6.3	.06[c]	<.001[b]
Medication use, decrease or increase (%)				
Opiate	↓23.2	↓44.3	.02[b]	<.001[b]
NSAIDs[d]	↓9.2	↓12.8	.08[c]	.06[c]
Acetaminophen	↑14.6	↑21.0	<.001[b]	<.001[b]
Gabapentin	↑4.6	↓76.3	.71	<.001[b]
Acceptable pain, decrease or increase (%)				
Knee at rest	↓31.3	↓20.3	.11	.25
Knee during activity	↓17.2	↑31.1	.40	.38
Knee at night	↓21.1	↑36.4	.21	.25
Experience with the executed recommended exercises—exactly enough, decrease or increase (%)	↓33.1	↓8.7	.02[b]	.67

[a]VAS: visual analog scale.

[b]Significant difference ($P<.05$).

[c]Trend ($.05<P<.10$).

[d]NSAIDs: nonsteroidal anti-inflammatory drugs.

Table 4. Findings in the PainCoach group, active PainCoach subgroup, and control group 1 month after surgery.

	1. PainCoach (n=38)	2. Active PainCoach (n=19)	3. Control (n=33)	*P* value (1 vs 3)	*P* value (2 vs 3)
VAS[a] pain, median (IQR[b])					
Knee at rest	11.5 (5.0-20.8)	11.5 (4.3-18.8)	10.0 (5.0-25.0)	.77	.53
Knee during activity	14.0 (7.0-28.8)	12.5 (9.3-26.3)	15.0 (8.0-35.0)	.49	.59
Knee at night	15.0 (7.0-33.0)	15.0 (5.0-33.0)	15.0 (7.0-27.8)	.79	.89
Acceptable pain, n (%)					
Knee at rest	31 (96.9)	16 (100.0)	28 (96.6)	>.99	>.99
Knee during activity	30 (93.8)	15 (93.8)	25 (86.2)	.41	.64
Knee at night	26 (81.3)	14 (87.5)	26 (89.7)	.48	>.99
KOOS-PS[c], mean (SD)	36.5 (10.5)	33.5 (8.4)	39.6 (9.8)	.24	.04[d]
OKS[e], mean (SD)	28.4 (8.4)	29.9 (9.1)	26.8 (6.2)	.42	.18
EQ-5D[f] descriptive system, median (IQR)	0.775 (0.693-0.843)	0.811 (0.775-0.857)	0.775 (0.651-0.811)	.34	.11
EQ VAS[g], median (IQR)	80.0 (70.0-90.0)	83.5 (70.0-90.0)	80.0 (65.5-89.5)	.56	.32

[a]VAS: visual analog scale.

[b]IQR: interquartile range.

[c]KOOS-PS: Knee Injury and Osteoarthritis Outcome Score–Physical Function Short-form.

[d]Significant difference (*P*<.05).

[e]OKS: Oxford Knee Score.

[f]EQ-5D: EuroQol-5 Dimensions.

[g]EQ VAS: EuroQol visual analog scale.

Other Pain Medication Use, Pain Acceptance, and Experience With Executed Recommended Exercises

In the PainCoach group, there was a statistically significant 14.6% increase in acetaminophen use (95% CI 8.2-21.3; *P*<.001) and no statistically significant differences in NSAID use and gabapentin use when compared with the findings in the control group during the first 2 weeks at home (Table 3). Overall pain medication use was below the advised maximum in both groups. Pain acceptance was 86.5% at rest, 86.5% during activity, and 79.4% at night in the PainCoach group and was 90.4% at rest, 88.6% during activity, and 83.0% at night in the control group, without statistically significant differences between the two groups. Regarding experience with executing recommended exercises, the PainCoach group had statistically significant 33.1% reduced experience with executing exactly enough exercises when compared with the findings in the control group (69.7% vs. 77.5%; 95% CI −52.0 to −6.7; *P*=.02) (Table 3). At 1 month after surgery, no statistically significant differences were found when comparing both groups (Table 4).

Adjusted analyses comparing the active PainCoach subgroup with the control group showed statistically significant 21.0% increased acetaminophen use in the active PainCoach subgroup (95% CI 12.6-30.0; *P*<.001) during the first 2 weeks at home. Additionally, the active PainCoach subgroup had statistically significant 76.3% decreased gabapentin use when compared with the findings in the control group (mean 0.1 [SD 0.3] tablets a day vs. 0.4 [SD 1.0] tablets a day; 95% CI −86.0 to −59.8; *P*<.001) (Table 3). In the active PainCoach subgroup, pain acceptance was 88.4% at rest, 90.9% during activity, and 87.4% at night. Regarding pain acceptance and experience with executing recommended exercises, no statistically significant differences were found between the active PainCoach subgroup and control group (Table 3). One month after surgery, the mean KOOS-PS was significantly lower in the active PainCoach subgroup (33.5 [SD 8.4]) than in the control group (39.6 [SD 9.8]) (*P*=.048) (Table 4).

PainCoach App Use

Among 28 patients who provided appropriate responses, 25 (89%) reported ease of app use, 22 (79%) found that the app added value, and 22 (79%) would recommend the app to friends and family. The PainCoach app was used 12 (IQR 4.5-22.0) times per patient on 7 (IQR 4.0-9.0) days at home. The number of patients with at least one entry in the PainCoach app ranged from 11 (30%) to 26 (70%) per day at home (Multimedia Appendix 2). The app was most frequently used between 9 and 10 am and mostly for advice on bearable pain.

Discussion

Principal Findings

This study aimed to determine the effects of an eHealth app, the PainCoach app, on pain control and opiate use in patients who underwent TKR during the first 2 weeks at home after surgery. The hypothesis was that the app would decrease pain and opiate use. As indicated by the main findings, there was no statistically significant difference in pain scores between the two groups and opiate use was significantly reduced by 23.2%

in the PainCoach group when compared with the finding in the control group. In the active PainCoach subgroup, however, pain during activity and at night significantly decreased 4.1 and 6.3 times faster, respectively, and opiate use significantly reduced by 44.3% when compared with the findings in the control group.

Overall, low pain scores and high levels of pain acceptance were found in this study. Only 21% (8/38) of patients in the PainCoach group and 30% (10/33) in the control group classified their pain as severe during one or more days at home. Other studies have stated that the most painful period after TKR surgery was the initial period at home, with 23%-30% of patients rating their average pain as severe [40,41]. Aside from the use of modern LIA techniques and a step-wise pain management protocol postoperatively, a possible explanation for the reported low pain and high acceptance scores in this study could be the guidance program that was provided to all patients who underwent TKR in Kliniek ViaSana. As less anxiety is associated with lower pain scores [14,19], the guidance provided might have resulted in less anxiety and therefore lower pain scores. The reported overall low pain scores also probably explain why no difference in pain scores was found between the PainCoach group and control group. Although overall pain scores were low, active use of the PainCoach app resulted in even lower pain scores during activity and at night when compared with the findings in the control group. These findings are in line with the results of a previous study showing that pain decreased by 0.7 points on a scale from 0 to 10 in patients with OA after online “pain coping skills” training [29]. Others have stated that 80% of interactive information is remembered compared with 20% of auditory information and 40% of read information [30,42,43]. As the PainCoach app is an interactive tool, it is logical that active use will result in better use of the pain management strategies provided and subsequently lower pain scores.

Opiate addiction caused 74 deaths in the Netherlands in 2016, and this number is increasing each year [44]. Using the PainCoach app, opiate use reduced by 23.2%, and active PainCoach app use resulted in a further reduction (44.3%). Because of a lack of standardized opiate prescribing protocols in orthopedic surgery, it is difficult to compare the reported amount of opiate use in this study with that in other studies. In one available study, a daily average morphine dose at discharge of 155 (SD 63) mg was prescribed to patients who underwent TKR, which would be the equivalent of 11 tablets per day of the opiate used in this study (oxycodon, 5 mg per tablet) and is far above the average use of 0.4 opiate tablets per day in this study [45]. The low preoperative opiate use of patients in this study might have contributed to the low opiate use after surgery, as preoperative opiate use is a strong predictor for prolonged opiate use after TKR [42,46,47]. With lower opiate use, acetaminophen use was higher, with a 14.6% increase in the PainCoach group and 21.0% increase in the active PainCoach subgroup. It can be concluded that because of the advice provided by the PainCoach app, opiate use was substituted by acetaminophen use. Opiate use was only advised in the presence of severe enough reported pain in the app. Therefore, it is concluded that the app helps to reduce the risk of the adverse effects of opiate use [48,49].

A shorter hospital stay is associated with a higher burden among patients, who need to take responsibility for aftercare shortly after surgery. Recent studies have shown that patients feel uncertain and left alone after discharge, which could increase anxiety and affect their pain coping and subsequent management [50,51]. Patients might need more individualized guidance, and the PainCoach app was developed to satisfy this need. The app scored high on usability, likelihood of being recommended to others, and added value. The results of this study show that the PainCoach app is a successful pain management tool, and its active use is recommended for the best effects on pain and opiate use.

To our knowledge, this is the first randomized controlled trial to examine the effects of eHealth with regard to controlling pain and reducing opiate use after TKR. The strengths of this study are that the actual amount of app use was measured and because of the unique download codes adopted, it was not possible for the control group to use the PainCoach app. The shortcomings are that the additional analysis was underpowered and the cost-effectiveness of the PainCoach app was not investigated. Furthermore, as there is no short validated questionnaire in Dutch for measuring pain acceptance, an expert group decided to assess pain acceptance using happy and sad smileys as the best alternative. In the population of this study, opiate use was already low. The app might have a much stronger effect in patient populations where preoperative opiate use is much higher. It is questionable if the PainCoach app is effective in the overall TKR population, as this study investigated the effects in patients having ASA I-II and BMI ≤35, which represent around 80% of the total TKR population [52,53]. Future research should focus on a larger sample size of the total TKR population, determination of the cost-effectiveness of the app, and use of the app in populations that have much higher preoperative opiate use.

Conclusions

The use of the PainCoach app contributes to reduced opiate use in the initial period at home after TKR. Active use of this app leads to further reduction in opiate use and improved pain control.

Acknowledgments

We thank the nurses for their help with executing this study, Klaartje Pijnappels for her assistance with data collection, and all patients who participated in this study.

Authors' Contributions

YP was responsible for designing the study, data collection, and study coordination; MP performed the statistical analysis supervised by YP; and YP, MP, AS, and MB were involved in drafting and revising the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Pain management protocol.
[PDF File (Adobe PDF File), 409 KB - mhealth_v8i3e16415_app1.pdf]

Multimedia Appendix 2
Content and use of the PainCoach app.
[PDF File (Adobe PDF File), 302 KB - mhealth_v8i3e16415_app2.pdf]

Multimedia Appendix 3
CONSORT - EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2971 KB - mhealth_v8i3e16415_app3.pdf]

References

1. Jüni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. Best Pract Res Clin Rheumatol 2006 Aug;20(4):721-740. [doi: 10.1016/j.berh.2006.05.002] [Medline: 16979535]
2. de Beer JD, Winemaker MJ, Donnelly GA, Miceli PC, Reiz JL, Harsanyi Z, et al. Efficacy and safety of controlled-release oxycodone and standard therapies for postoperative pain after knee or hip replacement. Can J Surg 2005 Aug;48(4):277-283 [FREE Full text] [Medline: 16149361]
3. Lin P. An evaluation of the effectiveness of relaxation therapy for patients receiving joint replacement surgery. J Clin Nurs 2012 Mar;21(5-6):601-608. [doi: 10.1111/j.1365-2702.2010.03406.x] [Medline: 21306457]
4. Dalury DF. A state-of-the-art pain protocol for total knee replacement. Arthroplast Today 2016 Mar;2(1):23-25 [FREE Full text] [doi: 10.1016/j.artd.2016.01.004] [Medline: 28326393]
5. Kehlet H, Wilmore DW. Evidence-based surgical care and the evolution of fast-track surgery. Ann Surg 2008 Aug;248(2):189-198. [doi: 10.1097/SLA.0b013e31817f2c1a] [Medline: 18650627]
6. Wu CL, Rowlingson AJ, Partin AW, Kalish MA, Courpas GE, Walsh PC, et al. Correlation of postoperative pain to quality of recovery in the immediate postoperative period. Reg Anesth Pain Med 2005;30(6):516-522 [FREE Full text] [doi: 10.1016/j.rapm.2005.07.190] [Medline: 16326335]
7. Trasolini NA, McKnight BM, Dorr LD. The Opioid Crisis and the Orthopedic Surgeon. J Arthroplasty 2018 Nov;33(11):3379-3382.e1. [doi: 10.1016/j.arth.2018.07.002] [Medline: 30075877]
8. Dowsey MM, Nikpour M, Choong PF. Outcomes following large joint arthroplasty: does socio-economic status matter? BMC Musculoskelet Disord 2014 May 06;15:148 [FREE Full text] [doi: 10.1186/1471-2474-15-148] [Medline: 24885773]
9. Fillingim RB, King CD, Ribeiro-Dasilva MC, Rahim-Williams B, Riley JL. Sex, gender, and pain: a review of recent clinical and experimental findings. J Pain 2009 May;10(5):447-485 [FREE Full text] [doi: 10.1016/j.jpain.2008.12.001] [Medline: 19411059]
10. Jiang Y, Sanchez-Santos MT, Judge AD, Murray DW, Arden NK. Predictors of Patient-Reported Pain and Functional Outcomes Over 10 Years After Primary Total Knee Arthroplasty: A Prospective Cohort Study. J Arthroplasty 2017 Jan;32(1):92-100.e2 [FREE Full text] [doi: 10.1016/j.arth.2016.06.009] [Medline: 27444848]
11. Racine M, Tousignant-Laflamme Y, Kloda LA, Dion D, Dupuis G, Choinière M. A systematic literature review of 10 years of research on sex/gender and experimental pain perception - part 1: are there really differences between women and men? Pain 2012 Mar;153(3):602-618. [doi: 10.1016/j.pain.2011.11.025] [Medline: 22192712]
12. Sullivan M, Tanzer M, Stanish W, Fallaha M, Keefe FJ, Simmonds M, et al. Psychological determinants of problematic outcomes following Total Knee Arthroplasty. Pain 2009 May;143(1-2):123-129. [doi: 10.1016/j.pain.2009.02.011] [Medline: 19304392]
13. Escobar A, Quintana JM, Bilbao A, Azkárate J, Güenaga JI, Arenaza JC, et al. Effect of patient characteristics on reported outcomes after total knee replacement. Rheumatology (Oxford) 2007 Jan;46(1):112-119. [doi: 10.1093/rheumatology/kel184] [Medline: 16735451]
14. Blackburn J, Qureshi A, Amirfeyz R, Bannister G. Does preoperative anxiety and depression predict satisfaction after total knee replacement? Knee 2012 Oct;19(5):522-524. [doi: 10.1016/j.knee.2011.07.008] [Medline: 21846588]

15. Edwards RR, Haythornthwaite JA, Smith MT, Klick B, Katz JN. Catastrophizing and depressive symptoms as prospective predictors of outcomes following total knee replacement. Pain Res Manag 2009;14(4):307-311 [FREE Full text] [doi: 10.1155/2009/273783] [Medline: 19714271]
16. Riddle DL, Wade JB, Jiranek WA, Kong X. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. Clin Orthop Relat Res 2010 Mar;468(3):798-806 [FREE Full text] [doi: 10.1007/s11999-009-0963-y] [Medline: 19585177]
17. Vissers MM, Bussmann JB, Verhaar JA, Busschbach JJ, Bierma-Zeinstra SM, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. Semin Arthritis Rheum 2012 Feb;41(4):576-588. [doi: 10.1016/j.semarthrit.2011.07.003] [Medline: 22035624]
18. Lingard EA, Riddle DL. Impact of psychological distress on pain and function following knee arthroplasty. J Bone Joint Surg Am 2007 Jun;89(6):1161-1169. [doi: 10.2106/JBJS.F.00914] [Medline: 17545417]
19. Brander VA, Stulberg SD, Adams AD, Harden RN, Bruehl S, Stanos SP, et al. Predicting total knee replacement pain: a prospective, observational study. Clin Orthop Relat Res 2003 Nov(416):27-36. [doi: 10.1097/01.blo.0000092983.12414.e9] [Medline: 14646737]
20. De Luca ML, Ciccarello M, Martorana M, Infantino D, Letizia Mauro G, Bonarelli S, et al. Pain monitoring and management in a rehabilitation setting after total joint replacement. Medicine (Baltimore) 2018 Oct;97(40):e12484 [FREE Full text] [doi: 10.1097/MD.0000000000012484] [Medline: 30290604]
21. Niemeläinen M, Kalliovalkama J, Aho AJ, Moilanen T, Eskelinen A. Single periarticular local infiltration analgesia reduces opiate consumption until 48 hours after total knee arthroplasty. A randomized placebo-controlled trial involving 56 patients. Acta Orthop 2014 Dec;85(6):614-619 [FREE Full text] [doi: 10.3109/17453674.2014.961399] [Medline: 25238439]
22. Reece AS, Hulse GK. Impact of lifetime opioid exposure on arterial stiffness and vascular age: cross-sectional and longitudinal studies in men and women. BMJ Open 2014 Jun 02;4(6):e004521 [FREE Full text] [doi: 10.1136/bmjopen-2013-004521] [Medline: 24889849]
23. Schoenfeld AJ, Jiang W, Chaudhary MA, Scully RE, Koehlmoos T, Haider AH. Sustained Prescription Opioid Use Among Previously Opioid-Naive Patients Insured Through TRICARE (2006-2014). JAMA Surg 2017 Dec 01;152(12):1175-1176 [FREE Full text] [doi: 10.1001/jamasurg.2017.2628] [Medline: 28813584]
24. Currie M, Philip LJ, Roberts A. Attitudes towards the use and acceptance of eHealth technologies: a case study of older adults living with chronic pain and implications for rural healthcare. BMC Health Serv Res 2015 Apr 16;15:162 [FREE Full text] [doi: 10.1186/s12913-015-0825-0] [Medline: 25888988]
25. Hochstenbach LM, Zwakhalen SM, Courtens AM, van Kleef M, de Witte LP. Feasibility of a mobile and web-based intervention to support self-management in outpatients with cancer pain. Eur J Oncol Nurs 2016 Aug;23:97-105 [FREE Full text] [doi: 10.1016/j.ejon.2016.03.009] [Medline: 27456381]
26. Jibb LA, Stevens BJ, Nathan PC, Seto E, Cafazzo JA, Johnston DL, et al. Implementation and preliminary effectiveness of a real-time pain management smartphone app for adolescents with cancer: A multicenter pilot clinical study. Pediatr Blood Cancer 2017 Oct;64(10). [doi: 10.1002/pbc.26554] [Medline: 28423223]
27. Machado GC, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: A systematic review. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1098-1109. [doi: 10.1016/j.berh.2017.04.002] [Medline: 29103552]
28. Naylor MR, Naud S, Keefe FJ, Helzer JE. Therapeutic Interactive Voice Response (TIVR) to reduce analgesic medication use for chronic pain management. J Pain 2010 Dec;11(12):1410-1419 [FREE Full text] [doi: 10.1016/j.jpain.2010.03.019] [Medline: 20620119]
29. Rini C, Porter LS, Somers TJ, McKee DC, DeVellis RF, Smith M, et al. Automated Internet-based pain coping skills training to manage osteoarthritis pain: a randomized controlled trial. Pain 2015 May;156(5):837-848 [FREE Full text] [doi: 10.1097/j.pain.0000000000000121] [Medline: 25734997]
30. Sandberg EH, Sharma R, Sandberg WS. Deficits in retention for verbally presented medical information. Anesthesiology 2012 Oct;117(4):772-779. [doi: 10.1097/ALN.0b013e31826a4b02] [Medline: 22902965]
31. Zickuhr K, Madden M. Pew Research Center's Internet & American Life Project. 2012. Older adults and internet use URL: https://www.sainetz.at/dokumente/studien/Older_adults_and_internet_use_2012.pdf [accessed 2018-02-08]
32. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983 Sep;17(1):45-56. [doi: 10.1016/0304-3959(83)90126-4] [Medline: 6226917]
33. Delgado DA, Lambert BS, Boutris N, McCulloch PC, Robbins AB, Moreno MR, et al. Validation of Digital Visual Analog Scale Pain Scoring With a Traditional Paper-based Visual Analog Scale in Adults. J Am Acad Orthop Surg Glob Res Rev 2018 Mar;2(3):e088 [FREE Full text] [doi: 10.5435/JAAOSGlobal-D-17-00088] [Medline: 30211382]
34. Chiu LY, Sun T, Ree R, Dunsmuir D, Dotto A, Ansermino JM, et al. The evaluation of smartphone versions of the visual analogue scale and numeric rating scale as postoperative pain assessment tools: a prospective randomized trial. Can J Anaesth 2019 Jun;66(6):706-715. [doi: 10.1007/s12630-019-01324-9] [Medline: 30796700]
35. Jamison RN, Gracely RH, Raymond SA, Levine JG, Marino B, Herrmann TJ, et al. Comparative study of electronic vs. paper VAS ratings: a randomized, crossover trial using healthy volunteers. Pain 2002;99(1):341-347. [doi: 10.1016/s0304-3959(02)00178-1]

36. Perruccio AV, Stefan LL, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS) - an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008 May;16(5):542-550 [FREE Full text] [doi: 10.1016/j.joca.2007.12.014] [Medline: 18294869]
37. Haverkamp D, Breugem SJ, Sierevelt IN, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the Oxford 12-item knee questionnaire for knee arthroplasty. Acta Orthop 2005 Jun;76(3):347-352. [Medline: 16156462]
38. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
39. Stomp-van den Berg SG, Vlaeyen JW, Ter Kuile MM, Spinhoven P, van Breukelen G, Kole-Snijders AM. Pijn Coping en Cognitie Lijst (PCCL). Meetinstrumenten chronische pijn deel 2 URL: https://docplayer.nl/13110836-Meetinstrumenten-chronische-pijn-deel-2-pijn-coping-en-cognitie-lijst-pccl.html [accessed 2018-02-08]
40. Chan EY, Blyth FM, Nairn L, Fransen M. Acute postoperative pain following hospital discharge after total knee arthroplasty. Osteoarthritis Cartilage 2013 Sep;21(9):1257-1263 [FREE Full text] [doi: 10.1016/j.joca.2013.06.011] [Medline: 23973139]
41. Aasvang EK, Luna IE, Kehlet H. Challenges in postdischarge function and recovery: the case of fast-track hip and knee arthroplasty. Br J Anaesth 2015 Dec;115(6):861-866 [FREE Full text] [doi: 10.1093/bja/aev257] [Medline: 26209853]
42. Bol N, Smets EM, Rutgers MM, Burgers JA, de Haes HC, Loos EF, et al. Do videos improve website satisfaction and recall of online cancer-related information in older lung cancer patients? Patient Educ Couns 2013 Sep;92(3):404-412. [doi: 10.1016/j.pec.2013.06.004] [Medline: 23820196]
43. Selinger CP, Carbery I, Warren V, Rehman AF, Williams CJ, Mumtaz S, et al. The relationship between different information sources and disease-related patient knowledge and anxiety in patients with inflammatory bowel disease. Aliment Pharmacol Ther 2017 Dec;45(1):63-74. [doi: 10.1111/apt.13831] [Medline: 27778366]
44. van Laar MW, van Gestel B, Cruts AA, van der Pol PM, Ketelaars AP, Beenakkers EM, et al. Nationale Drug Monitor - Jaarbericht 2017. Utrecht: Trimbos instituut; 2018.
45. Hernandez NM, Parry JA, Taunton MJ. Patients at Risk: Large Opioid Prescriptions After Total Knee Arthroplasty. J Arthroplasty 2017 Aug;32(8):2395-2398. [doi: 10.1016/j.arth.2017.02.060] [Medline: 28392133]
46. Bedard NA, DeMik DE, Dowdle SB, Callaghan JJ. Trends and risk factors for prolonged opioid use after unicompartmental knee arthroplasty. Bone Joint J 2018 Jan;100-B(1 Supple A):62-67 [FREE Full text] [doi: 10.1302/0301-620X.100B1.BJJ-2017-0547.R1] [Medline: 29292342]
47. Franklin PD, Karbassi JA, Li W, Yang W, Ayers DC. Reduction in narcotic use after primary total knee arthroplasty and association with patient pain relief and satisfaction. J Arthroplasty 2010 Sep;25(6 Suppl):12-16. [doi: 10.1016/j.arth.2010.05.003] [Medline: 20580191]
48. Sinatra RS, Jahr JS, Reynolds LW, Viscusi ER, Groudine SB, Payen-Champenois C. Efficacy and safety of single and repeated administration of 1 gram intravenous acetaminophen injection (paracetamol) for pain management after major orthopedic surgery. Anesthesiology 2005 Apr;102(4):822-831. [doi: 10.1097/00000542-200504000-00019] [Medline: 15791113]
49. Yeh Y, Reddy P. Clinical and economic evidence for intravenous acetaminophen. Pharmacotherapy 2012 Jun;32(6):559-579. [doi: 10.1002/j.1875-9114.2011.01085.x] [Medline: 22570116]
50. Specht K, Agerskov H, Kjaersgaard-Andersen P, Jester R, Pedersen BD. Patients' experiences during the first 12 weeks after discharge in fast-track hip and knee arthroplasty - a qualitative study. Int J Orthop Trauma Nurs 2018 Nov;31:13-19. [doi: 10.1016/j.ijotn.2018.08.002] [Medline: 30297138]
51. Sjøveian AK, Leegaard M. Hip and knee arthroplasty - patient's experiences of pain and rehabilitation after discharge from hospital. Int J Orthop Trauma Nurs 2017 Nov;27:28-35. [doi: 10.1016/j.ijotn.2017.07.001] [Medline: 28781113]
52. LROI. LROI Rapportage 2017. Distribution of ASA score of patients who underwent a primary knee arthroplasty per hospital in The Netherlands in 2016 URL: http://www.lroi-rapportage.nl/knee-primary-knee-arthroplasty-demographics-practice-variation-asa-score [accessed 2019-12-09]
53. LROI. LROI Rapportage 2017. Distribution of BMI of patients who underwent a primary knee arthroplasty per hospital in The Netherlands in 2016 URL: http://www.lroi-rapportage.nl/knee-primary-knee-arthroplasty-demographics-practice-variation-body-mass-index [accessed 2019-12-09]

Abbreviations

ASA: American Society of Anesthesiologists
BMI: body mass index
eHealth: electronic health
EQ: EuroQol
EQ-5D: EuroQol-5 Dimensions
EQ-5D-3L: EuroQol-5 Dimensions 3-level version
EQ VAS: EuroQol visual analog scale
IQR: interquartile range
KOOS-PS: Knee Injury and Osteoarthritis Outcome Score–Physical Function Short-form

LIA: local infiltration anesthesia
NSAIDs: nonsteroidal anti-inflammatory drugs
TKR: total knee replacement
VAS: visual analog scale

Edited by G Eysenbach; submitted 26.09.19; peer-reviewed by B Thompson, T Janmohamed; comments to author 12.11.19; revised version received 10.12.19; accepted 21.12.19; published 13.03.20.

Please cite as:
Pronk Y, Peters MCWM, Sheombar A, Brinkman JM
Effectiveness of a Mobile eHealth App in Guiding Patients in Pain Control and Opiate Use After Total Knee Replacement: Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(3):e16415
URL: http://mhealth.jmir.org/2020/3/e16415/
doi:10.2196/16415
PMID:32167483

Original Paper

Performance Evaluation of an Information Technology Intervention Regarding Charging for Inpatient Medical Materials at a Regional Teaching Hospital in Taiwan: Empirical Study

Min-Chi Liao[1,2], MBA; I-Chun Lin[1], PhD

[1]Department of Industrial Engineering and Management, National Yunlin University of Science and Technology, Yunlin, Taiwan

[2]Department of Nursing, National Taiwan University Hospital Yunlin Branch, Yunlin, Taiwan

Corresponding Author:
I-Chun Lin, PhD
Department of Industrial Engineering and Management
National Yunlin University of Science and Technology
Yunlin
Taiwan
Phone: 886 05 5342601 ext 5126
Email: ichunlin@yuntech.edu.tw

Abstract

Background: The process of manually recording the consumption of medical materials can be time consuming and prone to omission owing to its detailed and complicated nature. Implementing an information system will better improve work performance.

Objective: The Information System Success Model was adopted as the theoretical foundation. The opinions of nursing staff were collected to verify the impact of the system intervention on their work performance.

Methods: This cross-sectional study was conducted at a regional teaching hospital. Nursing staff were invited to participate in the field survey. A total of 296 questionnaires were collected, and of these, 284 (95.9%) were valid and returned.

Results: The key findings showed that two critical factors ("subjective norm" and "system quality") had significant positive effects (both $P<.001$) on user satisfaction ($R^2=0.709$). The path of "service quality" to "user satisfaction" showed marginal significance ($P=.08$) under the 92% CI. Finally, the explanatory power of the model reached 68.9%.

Conclusions: Support from the top management, appointment of a nurse supervisor as the change agent, recruitment of seed members to establish a pioneer team, and promotion of the system through the influence of opinion leaders in small groups were critical success factors needed for implementing the system in the case hospital. The target system was proven to be able to improve work performance, and the time saved could be further used for patient care, thereby increasing the value of nursing work. The positive experiences gained from this study could lay the foundation for the further promotion of the new system, and this is for future studies to replicate. The example of the successful experience of the case hospital could also serve as a reference for other hospitals in developing countries like Taiwan with regard to the promotion of nursing informatization.

(JMIR Mhealth Uhealth 2020;8(3):e16381) doi:10.2196/16381

KEYWORDS

Information System Success Model; information technology intervention; charging; medical materials; work performance

Introduction

Background

Computers having high performance are utilized in many industries to increase competitiveness and quality. Informatization has become one of the critical factors determining the success of a company. Information technology (IT) has been heavily introduced in many companies, and the highly complex medical industry is no exception. The National Health Insurance (NHI) program implemented in Taiwan in 1995 not only changed the existing medical ecology, but also affected the development of hospital information systems (HISs). Owing to the impact of changes in the medical payment system and the demand for improved medical quality, medical institutions are facing an enormous challenge. According to the provisions of the NHI Administration, medical expenses must be declared electronically. To comply with the health insurance declaration work, medical institutions had to immediately

computerize their operations at that time. In the early days of computerization, the major objective of the HISs established by medical institutions was to focus on collecting health insurance declaration data. Owing to the subsequent increased reliance on IT for daily operations and other demands, such as compliance with hospital evaluation and improvement of medical service quality, other relevant functions were gradually added to the HIS [1]. The NHI system now covers 99.6% of Taiwan's population, and it has service contracts with 93% of the country's hospitals and clinics [2]. Health insurance declaration is the hospital's primary income source, although the hospital's primary revenue comes from diagnostic and treatment services provided by physicians. Therefore, the initial phase of hospital informatization in Taiwan mainly focused on outpatient clinics and physicians, and computerization and informatization for inpatients and nursing care were only gradually implemented afterwards.

IT and network communications are widely used in various industries, and the communications as such have led to different levels of digital divide owing to differences in resource allocation, learning environments, and even urban-rural gaps [3]. Research in this field has shown that digital divide among hospitals mostly arises from different geographical locations and hospital levels [4]. Meanwhile, female gender and age are two of the major factors causing digital divide [5]. The study conducted by Venkatesh et al in 2003 [6] demonstrated that female gender and age are factors that highly influence user acceptance of new information systems (ISs). Nursing staff are mostly female, and they are disadvantaged regarding adaptation to digital technology because of their low exposure. Thus, they are more likely to encounter problem-solving difficulties when working with ISs [7]. The case hospital is located in an agricultural city in central Taiwan. It is a regional teaching hospital with approximately 941 beds and is a typical rural hospital in Taiwan. Compared with other metropolitan hospitals, the adoption and maturity of IT are comparatively low, and the information literacy of the nursing staff is relatively inadequate. The implementation of informatization in the inpatient nursing department of the case hospital started recently, and prior to this, most procedures were manually operated. For example, regarding charging for inpatient medical materials in the past, the consumption items were manually recorded on charging sheets by nurses after performing various treatments for patients, and the sheets were sent to clerks for manual charging and verification. In addition, nursing work showed several characteristics, including 24-hour continuity of care, a three-shift system, and a high patient-to-nurse ratio. In the case hospital, each nurse needs to take care of 8 to 10 patients in a typical day shift, 13 to 17 patients in a typical night shift, and 16 to 20 patients in a typical graveyard shift. Additionally, the nursing staff has a heavy workload, and nursing work is cumbersome. Thus, manual operations are prone to errors. There are often many disturbances in the ward. For example, when there is an emergency involving an inpatient, the nurse must drop the work at hand and immediately provide emergency care. This results in sudden interruption of the nursing work at hand. Furthermore, there are many treatment items to be managed, and the charging work is cumbersome. With this in mind, these factors may easily cause charging mistakes (eg, missed records, incorrect records, etc). Moreover, manual operations lack efficiency, and this negatively impacts hospital income.

Charging refers to the information included in the record of all financial transactions. Hence, the analysis and processing of charges can serve as a reference for decisions made by managers. To be specific, correct charging can not only reduce and prevent the risks of using expired medical materials, but also scrutinize medical shortage. Patient safety can be ensured, and the hospital costs can be reduced [8,9]. Correct charging and its analysis can facilitate hospital management and decision-making processes, as they will enable hospitals to understand the immediate need or demand for medical materials in each department, construct performance indicators [10,11], develop profit models, and plan hospital development policies [12,13]. Charging is based on the concept of structuralization, wherein an intervention IS can hopefully improve the efficiency of procedures and reduce workload. Because of the construct of structuralization and informatization of charging, numerous benefits can be introduced in hospitals. It can also simplify the charging procedure and provide flexibility for the use of the charging information effectively and efficiently. Work efficiency will therefore be increased. The same is the case for timely and accurate transmission of information [14]. Liu et al [15] aimed to improve the problem of missed charging information by improving the accuracy of wound care declaration in the emergency department and lowering the error rate of wound image data archiving. An app into which image data could be uploaded was first introduced in the emergency room. The data were transmitted to the electronic medical record (EMR) system for automatic archiving. This further solved the problem of omission of file uploading by emergency nurses owing to their busy schedule or work interruption. It also helped reduce declaration errors due to inaccurate archiving. Apart from reducing the workload of emergency nurses, this approach increases the accuracy of health insurance declaration for wound care and reduces revenue loss. As nurses are requested to provide patients with frontline care 24 hours a day, they are most familiar with the use and consumption of medical materials. The introduction of IT in nursing work would enable nurses to complete their duties more efficiently [16,17]. Moreover, an IS could allow instantaneous, cross-temporal, and cross-departmental integration of information. This integration would allow effective cross-departmental communication and co-ordination to occur. Thereby, the overall operational performance of the hospital would be enhanced to a great extent [18].

In view of the high importance of accurate charging and the potential deficiencies with manual charging in the past, the case hospital implemented an inpatient charging system to improve efficiency and effectiveness. However, in the literature, there is a lack of reports discussing whether ward nurses can really meet their needs at work with the use of hospital charging systems. Without effective user feedback for the management, user work efficiency may be negatively affected, which, in turn, could affect the performance of the organization. Therefore, understanding the factors impacting both user satisfaction and the IS will be of great research value for improving the effectiveness of IS implementation and management

performance [19,20]. The purpose of this research was to adopt the DeLone and McLean 2003 IS Success Model [21]. To that end, "subjective norm" was incorporated as a variable and "work performance" was incorporated as a dependent variable to better look into and understand the factors impacting nurses' satisfaction with the charging system and their work performance. Meanwhile, it is hoped that this study will help to improve the evaluation of the current status of system implementation in the case hospital, which could serve as a reference for subsequent system optimization. The same is the case for other rural and regional hospitals in Taiwan with backgrounds similar to those of the case hospital and hospitals located in other developing countries with relevant experience in relation to the introduction of the IS provided. More importantly, the results of this study will hopefully make valuable contributions to the completeness of relevant research on nursing informatization.

Literature Review

Prior Efforts of Applying the DeLone and McLean Information System Success Model to Measure Health Care Information Technology Success

The concept of the IS Success Model is based on an IS theory, which is established to provide a comprehensive understanding of IS success by identifying, describing, and explaining the relationships among the most critical six dimensions of success through which ISs are commonly evaluated. The IS success model has been widely cited in thousands of scientific papers, and this model is considered one of the most influential theories in contemporary IS research. Initial development of the theory was proposed by DeLone and McLean in 1992 [22]. From more than 180 relevant articles published on IS performance, they deduced the following six dimensions: system quality, information quality, system use, user satisfaction, individual impact, and organization impact. Although this model provided the key factors for consideration in academia with regard to IS success, some subsequent scholars, such as Seddon and Kiew [23], Seddon [20], and Pitt et al [24] raised some questions and proposals. In response to this, Pitt et al [24] pointed out that the model was too product oriented and that IT departments provide not only products but also services. Seddon [20] believed that the causal relationship of the model could rather confuse researchers. Therefore, DeLone and McLean [21] further refined the model a decade later in response to feedback received from other scholars working in the IS discipline [21,22]. The revised IS Success Model first included "services quality," and "use" was appropriately modified to "intention to use," with an explanation that "intention to use" was an attitude, whereas "use" was a behavior. Meanwhile, "use" and "user satisfaction" were posited to be interrelated, that is, "use" would lead to "user satisfaction," whereas "user satisfaction" would indirectly affect "use" through "intention to use." In doing so, a causal bidirectional relationship between "use" and "user satisfaction" existed. Finally, to adapt to the e-commerce environment, "net benefit" was used to indicate whether the overall result of using the IS was positive or negative. Collectively, the core of the IS Success Model is to enhance user perception regarding functional aspects, information processing and output, and service quality of the system through three major quality dimensions in order to strengthen the intention to use and satisfaction of the system, thereby promoting work performance (eg, improve work effectiveness, reduce errors, save time, etc). Specifically, "system quality" refers to the completeness of the system itself, and the indicators measured include response time, functional utility, ease of operation, compliance with users' needs, flexibility, accuracy, reliability, accessibility, and system integration [21,25]. "Information quality" is related with the quality of output information from the system, and the indicators measured include validity of information, preciseness, immediacy, completeness, relevance, and ease of understanding [21,22,25,26]. Service quality is derived from the well-known SERVQUAL scale [27], and the indicators measured include the components of tangibility, reliability, responsiveness, assurance, and empathy. High quality of service provided by the staff or suppliers is also suggestive of the capability and good intention of the service providers [28]. Pitt et al [24] revised the SERVQUAL scale to measure user perception and evaluation regarding the assistance or service provided by IT departments or system suppliers. Scholars deeply believe that IT departments or system suppliers should not only build and maintain ISs, but also provide services, such as problem solving, education, and training resources, to users [21,29]. User satisfaction is defined as the evaluation of a user's response after using the output information, and it is an important indicator for measuring IS effectiveness. The indicators measured include system interface satisfaction, software satisfaction, information satisfaction, decision-making satisfaction, and overall satisfaction of the system [21,30].

A review on health care information technology (HIT)–related research has shown that scholars mainly use the IS Success Model as a theoretical foundation to understand the impact of HIT on user behavior, and they have achieved good explanatory power [31-34]. Hwang et al [29] explored the factors considered by physicians regarding the use of EMRs from the benefit perspective. Bossen et al [32] adopted the IS Success Model to evaluate the outcomes of implementing an electronic health record (EHR), which involves conducting a survey among physicians and nursing staff. They found that implementing an EHR enabled hospital staff to grasp patients' conditions in a timely manner. Hsieh and Su [33] extended the IS Success Model and explored the key success factors of EMRs (impacts of implementation of EMRs from the perspective of medical and health information managers by combining an updated version of the DeLone and McLean IS Success Model [21]). Both Huang et al [35] and Chang and Lin [34] modified the IS Success Model to best explore the impact of introducing an IS to the shift system on the performance and satisfaction of nursing staff.

Furthermore, in the IS discipline, "user satisfaction" and "system usage or use" are the most commonly used key factors to measure the effectiveness of system implementation [21,22]. Ives et al [19] suggested that the measurement of "use" can only serve as a surrogate indicator for system success under specific conditions. The authors defined user satisfaction as the relative value of the IS perceived by users, which was the sum of user perceptions of different aspects of the IS, the evaluation

responses, and the attitude toward the IS [19]. Seddon and Kiew [23] defined user satisfaction as either the pleasant feeling or unpleasant feeling after using the IS. DeLone and McLean [21] pointed out that the relationship between "use" and "performance" was not significant if the users adopted the system in "compulsory" or "involuntary" situations. Under specific circumstances when users were requested to use the system (a common situation was requests by managers to use or comply with the implementation of the informatization policy), "use" was not a suitable surrogate indicator of IS success. In such cases, "user satisfaction" is more suitable than "use" as a surrogate indicator for measuring system effectiveness.

Previous research in the field of HIT often treated "user satisfaction" as a key factor in the success of IS implementation [33,34,36,37]. In the present situation, the inpatient charging system was the hospital's policy to promote informatization of wards and nursing care, and ward nurses were forced to use the inpatient charging system. Therefore, the "use" dimension was discarded, whereas the "user satisfaction" dimension was retained as an intermediary variable.

Work performance refers to all behaviors or actions related to organization goals, and the behaviors or actions can be measured according to individual proficiency and the different levels of contribution to the organization goals. It is believed that work performance should cover efficiency, effectiveness, and productivity [38]. Kast and Rosenzweig [39] believed that performance should include the efficiency, effectiveness, and participation satisfaction of the employees within an organization. Huang et al [35] explored the impact of introducing an IS to the shift system on nursing work performance and satisfaction. Work performance was defined as "the contribution of individual effort by nurses to the organization's missions and the interactions with other members within a certain period of time that met the organization's expectation; the quantity of work completed; and the value and quality of work contribution." Their research findings showed that the satisfaction obtained by nursing staff from using the IS shift system improved work performance to some extent. This part of the results is in line with the findings of the study by Chang and Lin [34]. According to the aforementioned literature, this study used "work performance" as a dependent variable to replace the "net benefit" dimension. The aim of this study was to explore the impact of the introduction of the new inpatient charging system on the performance and satisfaction of nursing staff.

Social Norm and Prior Health Care Information Technology–Related Research

Fishbein and Ajzen [40] defined "subjective norm" as the concept of how an individual experiences the perceptions of others important to him or her. For instance, whether they think that the individual should perform a particular action. Hence, subjective norms (social influence factors) are considered a direct determinant of behavioral intention. Ajzen [41] mentioned that subjective norms are the perceived expectations from others that influence a user to have a particular behavior. Subjective norms refer to the rules, regulations, or instructions designed by society on which one should act. For any specific behavior, society has set or designed norms on how to perform that behavior. Subjective norms refer to the kinds of norms an individual follows owing to social pressure. Many previous IS studies have confirmed that a stronger perception of subjective norms affects users' behavior when using a new system. Thus, subjective norms are considered important factors affecting users' acceptance or rejection of a new system [42-45]. Garcia-Smith and Effken [46] proposed the Clinical IS Success Model by integrating the Technology Acceptance Model with the IS Success Model. Specifically, social influence was also incorporated into the model to investigate the key factors for successful IS implementation through the evaluation of multidimensional factors. Their study has become a reference model for evaluating the success of IS implementation during the introduction of nursing informatization.

The medical industry has a high degree of industry specificity. Medical service is characterized by the irreversibility of life. The nature of innovation is uncertain, and the consequences of adopting innovations are often difficult to predict. For these reasons, medical personnel tend to respond to the introduction of innovations with more conservative strategies, such as observing and learning from others' experiences and opinions. They tend to reduce uncertainty about innovation through information from peers and the related social system. The study by Kuo et al [47] showed that the pressure caused by subjective norms, together with incentive measures, effectively motivated physicians to comply with the promotion of the EMR system. Further investigations revealed that the support and execution abilities of senior managerial staff were the keys to promote EMRs in hospitals. The utilization of influence from managers and peers facilitated the launch of the new IS and improved work performance. Huang et al [48] also pointed out that the attitudes of public health care workers toward the use of public health ISs were greatly affected by subjective norms (including peers and superiors). With regard to patient care, members of the care team provide patients with continuous, accurate, and complete care services by collaborating with each other. Together, the care team solves the patients' problems and assists with their recovery. As a result, care services are inherently highly mission-dependent. The education and professional characteristics of nursing staff are related to performing nursing work by following physicians' orders and respecting norms, such as clinical guidelines. Nursing staff members are more compliant than other medical staff, and they will consider the opinions given by important people around them (eg, supervisors and peers) or important communities and members when performing nursing work. In addition, at the beginning of the introduction of the inpatient charging system in the case hospital, the head nurse at each nursing station assigned nurses with higher information literacy as seeds and set up special mission teams to assist the promotion of informatization of the charging system. Based on the above factors, this study also incorporated social norms into the research framework. Together with three quality dimensions, this study utilized multidimensional factors to discuss the satisfaction of ward nurses with the hospital charging system and its impact on work performance.

In summary, the introduction of a nursing-related IS could increase the work efficiency of nurses and thus promote patient safety [49-51]. This study validated the construct, with the aim of providing a more efficient working environment for nurses along with the additional safety and security needed for patients.

Methods

Ethical Approval

This study was approved by the Human Research Ethics Committee, National Taiwan University Hospital Yunlin Branch (Institutional/Independent Review Board number: NTUH1063703472).

Target System Implementation and Operation

The hospital where the study was conducted mainly co-operated with different departmental teams from the IT and nursing departments to introduce the new inpatient charging system designated to improve the way of charging for the use of medical materials. This innovation was to replace the manual procedure. After the introduction of the new system, the huge difference observed was that the procedures needed to charge for care services were simplified to a greater extent. All that the nurses had to do was to check the items shown on the interface by clicking on them. In doing so, the money charged for the use of medical materials was automatically transmitted to the charging system, and because of this, clerks, instead of nurses, were mainly held accountable for the cumbersome and tedious tasks related to managing accounting details and financial statements.

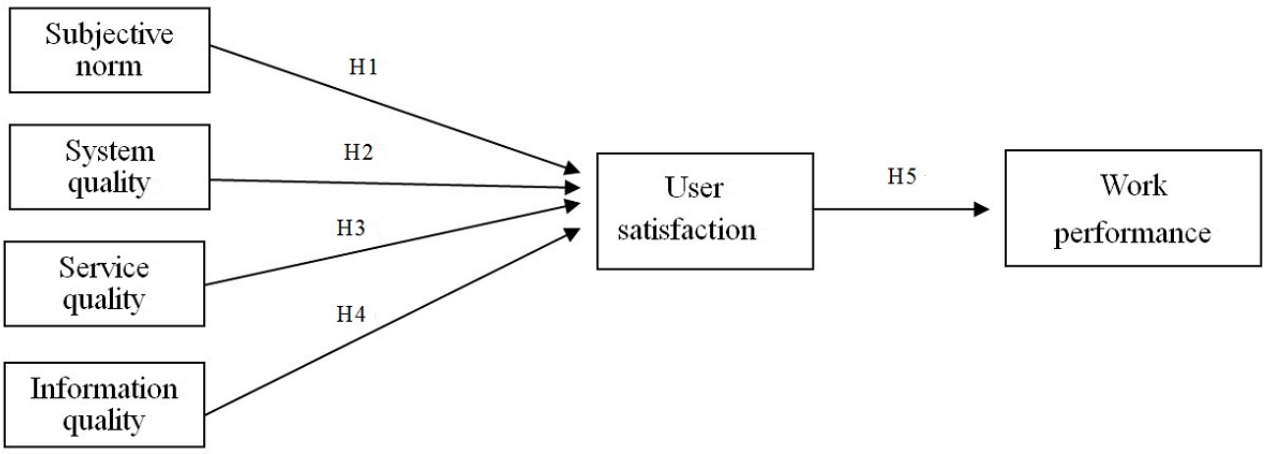
Research Model and Hypotheses

This is a cross-sectional study targeting a regional teaching hospital, which has adopted the e-hospital paperless policy and has implemented the charging system (as the target system mentioned below) for charging medical materials in the inpatient department. Six variables and five hypotheses (Textbox 1) were proposed in our model (Figure 1), including subjective norm, system quality, service quality, information quality, user satisfaction, and work performance. Our hypothetical model was empirically tested using data collected from a field survey. Regarding data collection, convenience sampling was used. Nursing staff who had used the system for at least 6 months or longer were invited to participate in the field survey. Participants who agreed to participate in this study filled out a self-reported questionnaire anonymously to protect their privacy. Thirty-nine items were generated according to our literature review. A 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree) was adopted in this study. The instrument (Multimedia Appendix 1) was reviewed and approved by two medical and nursing informatics scholars, three management IS scholars, and two nursing department managers. The experts all have more than 10 years of working experience. The items were finalized after modification according to expert suggestions. The scale's content validity was confirmed by professionals from the academic and industrial fields. A total of 296 questionnaires were administered, and of these, 284 were returned, resulting in a valid questionnaire return rate of 95.9%.

Textbox 1. The research hypotheses of this study.

H1: Subjective norm predicts nursing staff satisfaction with the target system.
H2: System quality predicts nursing staff satisfaction with the target system.
H3: Service quality predicts nursing staff satisfaction with the target system.
H4: Information quality predicts nursing staff satisfaction with the target system.
H5: The extent to which the work performance of nursing staff improves can be predicted through user satisfaction with the target system.

Figure 1. The hypothesized target system success model.



Data Analysis

Descriptive analysis and analysis of variance (ANOVA) were performed using IBM SPSS Statistics 20.0 (IBM Corp, Armonk, New York, USA) to enhance our understanding of the sample characteristics. Structural equation modeling was conducted using partial least squares path modeling in the software package Smart PLS version 3.0 (Smart PLS GmbH, Bönningstedt, Germany). Two items with loading lower than the recommended value were iteratively deleted from the model. The path coefficients for the trimmed model were therefore calculated and tested.

The reliability and validity of the measurement model were assessed by its psychometric properties. The psychometric properties of the model were assessed according to internal consistency and convergent and discriminant validity. For reflective indicators, internal consistency was measured according to composite reliability and Cronbach alpha [36], with a recommended acceptable value of 0.70 [37]. Convergent validity was measured according to the average variance extracted (AVE), and it was considered adequate when the AVE of each construct reached 0.50 [39]. Discriminant validity was considered the extent to which a variable is truly distinct from other variables [36]. It was considered acceptable when the square root of the AVE of each construct exceeded the correlation coefficient between the specific construct and others in the model. Discriminant validity was verified by factor loading and cross loading. The loading of an indicator on its assigned variable should be greater than its cross loading on all other variables. Moreover, a structural model is considered to include unobservable latent variables and the theoretical relationships among them [39]. It also suggests how well the theoretical model predicts the hypothesized paths or relationships.

Results

Samples

A total of 284 samples were selected, and they consisted of 29 seed members and 255 nonseed members. The analyses were performed as shown below.

Characteristics of the Participating Nursing Staff

The characteristics of the participating nursing staff are presented in Table 1. We found that the majority of the participants had university degrees (220/284, 77.5%), followed by associate degrees (54/284, 19.0%) and graduate degrees (10/284, 3.5%). In terms of age, the majority of the participants were between 21 and 35 years (241/284, 84.9%), followed by between 36 and 40 years (29/284, 10.2%) and between 41 and 50 years (11/284, 3.9%). With regard to nurse competency advancement, the majority of the participants (113/284, 39.8%) were at the N2 level. In terms of seniority, the majority of the participants (171/284, 60.2%) had less than 6 years of job experience. Among the 284 participants, 29 (10.2%) were serving as seed members, who would be promoting and assisting each unit to launch the target system. Similar to the national data [52] in Taiwan, 95.4% (271/284) of the survey respondents were female. Additionally, 85.6% (243/284) of the nursing staff members were under 36 years of age, 14.1% (40/284) were between 36 and 50 years of age, and only 0.4% (1/284) were over 60 years of age. Overall, our sample characteristics are similar to those of the population of nursing staff members across Taiwan.

Table 1. Characteristics of the participating nursing staff (n=284).

Variable	Value, n (%)
Gender	
Male	13 (4.6%)
Female	271 (95.4%)
Job title	
Nurse practitioner	21 (7.4%)
Nurse	251 (88.4%)
Head nurse	12 (4.2%)
Education	
Associate degree	54 (19.0%)
University degree	220 (77.5%)
Graduate degree	10 (3.5%)
Seniority, years of experience	
≤2	85 (29.9%)
3-5	86 (30.3%)
6-10	57 (20.1%)
≥11	56 (19.7%)
Age, years	
≤20	2 (0.7%)
21-25	112 (39.4%)
26-30	75 (26.4%)
31-35	54 (19.0%)
36-40	29 (10.2%)
41-45	5 (1.8%)
46-50	6 (2.1%)
51-60	0 (0.0%)
≥61	1 (0.4%)
Nurse competency advancement	
N	24 (8.5%)
N1	82 (28.9%)
N2	113 (39.8%)
N3	46 (16.2%)
N4	19 (6.7%)
Seed members or not	
Seed members	29 (10.2%)
Nonseed members	255 (89.8%)

Analysis of Variance

In order to enhance our understanding of the sample characteristics, variance analysis was performed by carrying out one-way ANOVA on “nurse competency advancement” (five levels), “seniority” (four intervals), and “whether users served as seed members for promoting nursing informatization” for the six research variables. The results indicated that there were insignificant differences for “nurse competency advancement” (P=.84 for “nurse competency advancement” to “user satisfaction” and P=.96 for “nurse competency advancement” to “work performance”) and “seniority” (P=.79 for “seniority” to “user satisfaction” and P=.84 for “seniority” to “work performance”).

However, regarding “whether users served as seed members for promoting nursing informatization,” significant differences

were found for all variables (*P*=.02) (Table 2). Further examination revealed that seed members had higher scores for the averages of the six variables as compared with nonseed members (Table 2). Seed members also evaluated the target system more favorably, showing higher scores in information quality, service quality, system quality, user satisfaction, and working performance.

Table 2. Results of analysis of variance and average scores according to whether users served as seed members for promoting nursing informatization.

Variable	*F* value	*P* value	Score, mean (SD)	
			Seed members (n=29)	Nonseed members (n=255)
Subjective norm	9.876	.002	4.3 (0.5)	3.9 (0.6)
System quality	9.263	.003	4.2 (0.6)	3.9 (0.6)
Service quality	12.949	<.001	4.2 (0.6)	3.7 (0.6)
Information quality	12.319	.001	4.3 (0.5)	3.9 (0.6)
User satisfaction	9.418	.002	4.2 (0.5)	3.9 (0.6)
Work performance	10.980	.001	4.1 (0.5)	3.8 (0.6)

Measurement Model Testing: Reliability and Validity of the Questionnaire

In terms of convergence validity and reliability, the factor-loading values of the dimensions were all greater than 0.6 (Table 3), the AVE values were all observed to be greater than 0.5 (Table 3), and the composite reliability and Cronbach α were all greater than .7, indicating good convergence validity and internal consistency overall. In addition, the Fornell-Larcker criterion of interconstruct correlations and cross-loading were used to confirm discriminant validity (Table 3). The diagonal shown in Table 3 was the square root of the AVE, and its minimum value (0.824) was higher than that of any other correlation coefficient in terms of all the other constructs involved in the test, except for user satisfaction to work performance (0.830). The results of the cross-loading analysis indicated that the scale had acceptable discriminant validity.

According to the ANOVA results, there were significant differences between users who were seed members (n=29) and those who were nonseed members (n=255). The data analyses were performed without the seed members (n=255) and with the seed members (total n=284) separately. This was to further explore whether other related possible outcomes were shown owing to the scale properties, research model, and hypotheses. However, the results indicated high similarity between these two groups. Therefore, we present the analysis results of the nonseed members (n=255).

Table 3. Scale properties (n=255).

Variable	AVE[a]	CR[b]	Cronbach α	Interconstruct correlations					
				IQ[c]	PF[d]	SEQ[e]	SN[f]	SYSQ[g]	SAT[h]
IQ	0.834	0.938	.900	0.913					
PF	0.727	0.955	.946	0.676	0.853				
SEQ	0.847	0.971	.964	0.545	0.613	0.920			
SN	0.884	0.958	.935	0.677	0.782	0.494	0.940		
SYSQ	0.678	0.950	.940	0.823	0.745	0.553	0.684	0.824	
SAT	0.957	0.985	.977	0.683	0.83	0.539	0.773	0.765	0.978

[a]AVE: average variance extracted.

[b]CR: composite reliability.

[c]IQ: information quality.

[d]PF: work performance.

[e]SEQ: service quality.

[f]SN: subjective norm.

[g]SYSQ: system quality.

[h]SAT: user satisfaction.

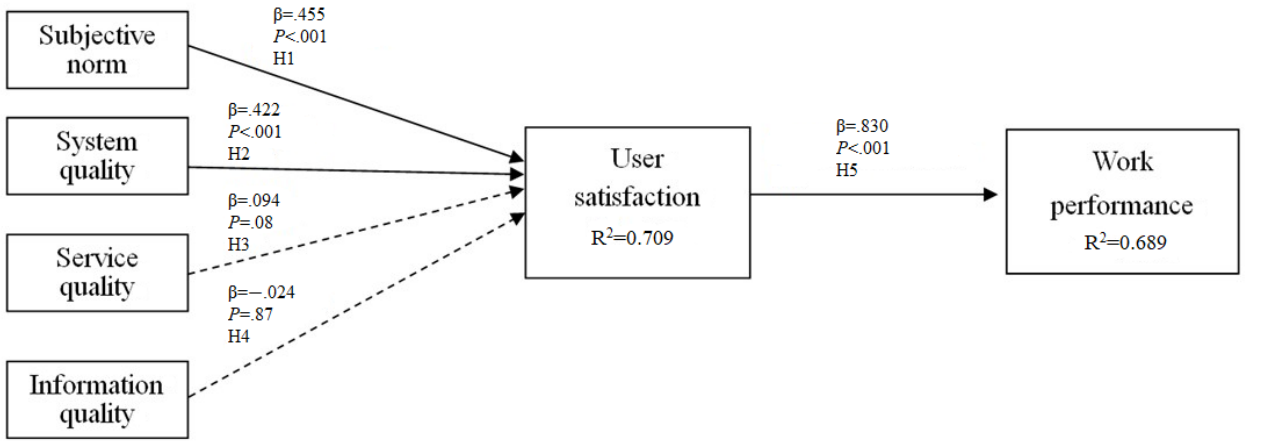
Path Analysis

The results of the path analysis indicated that three suggested research hypotheses (H1, H2, and H5) reached statistical significance (*P*<.001).

One path showed marginal significance (*P*=.08) for H3 under the 92% CI. The path from information quality to user satisfaction (H4) did not reach statistical significance (*P*=.87). The path coefficients were 0.455 (*P*<.001) for subjective norm to user satisfaction, 0.422 for system quality to user satisfaction,

0.094 for service quality to user satisfaction, and 0.830 for user satisfaction to work performance (Figure 2). The overall explanatory power of the model reached 68.9%.

Figure 2. Results of the research model and hypothesis validation.



Discussion

Principal Findings

The demographic information of the participants showed that there was a high proportion of female users (271/284, 95.4%) and that the majority of the participants were between 21 and 40 years of age (270/284, 95.1%). The real-world national data on the nursing population also show that most nurses are between 21 and 40 years of age (114,269/172,897, 66.1%) [52]. The threshold age for marriage is generally estimated to be around 40 years, because this is a suitable age for nurses to start a family. In doing so, they may choose to switch or change their own career or leave the current job owing to the need for child care after marriage or the inability to take turns working in shifts after parental leave. Furthermore, nursing work has the characteristics of work shifts, a three-shift system, and a high patient-to-nurse ratio. Thus, the work is physically challenging for most nursing staff as they grow older. For these aforementioned reasons, the majority of them are between 21 and 40 years of age, and the number of nurses aged over 41 years is gradually decreasing, as age is considered a great challenge if they continue to work as practice nurses.

According to the variance analysis, the nursing staff involved in this study showed relevant differences in their perceptions and evaluations of the target system in terms of all the research variables. Those nurses serving as seed members showed a significantly higher average score as compared with that for nonseed members (Table 2).

We believe that this was related to the fact that nurses who served as seed members were more familiar with the system than those who did not serve as seed members. Seed members were highly involved in the development, discussion, and promotion of the new charging system. With the experience of introducing the new IT charging system, seed members were more likely to facilitate the development of this new system. Their successful experiences could also be considered for replication in further studies regarding the introduction of a new IT system in another medical organization in the future.

The results obtained from the model validation indicated that both the subjective norm and system quality had a significant positive impact on user satisfaction (R^2=0.709, P<.001), and thus, these two major factors played decisive roles in the implementation of the target system. Meanwhile, this also demonstrated a positive influence on performance through user satisfaction (R^2=0.689). These results indicate that when a new system is introduced, the positive influence of opinion leaders and the enhancement of users' awareness and perception of the functional aspect of the system in the initial phase will provide users a chance to perceive the benefits of the new system intervention for their work. This will help increase user satisfaction, which will, in turn, make users more willing to use the system, leading to good work performance. Although the path of service quality to satisfaction exhibited marginal significance (92% CI) and the path coefficient was only 0.094, previous research has already demonstrated the importance of service quality. Therefore, the benefits of improving service quality for the improvement of user satisfaction should not be neglected, as they will, in turn, have positive effects on users' work performance.

In the case hospital, a chief nursing supervisor was assigned as the change agent when nursing informatization was promoted. In addition to promoting the new system, the change agent also served as a window for communication between the case hospital and the nursing informatization promotion team, as well as the IT department of the headquarter hospital. This dedicated agent also set up a pioneer team by recruiting colleagues with higher information literacy and personalities that were related to being more accepting of changes at each nursing station as seed members. Seed members were the first to receive education and training to familiarize themselves with the system operation. A total of 29 seed members provided successive assistance to promote the new system and remove barriers to system usage for colleagues within the unit. As the case hospital was located in a rural area and more than 95.4%

of the nurses were female, the information capability was relatively weak. This was in line with the study conducted by Lin and Lee [7], who pointed out that when most nursing staff are female, there could be a disadvantage with regard to adaptation to digital technology owing to their limited exposure. Thus, it is highly likely that they would face problem-solving challenges when using an IS. Therefore, the case hospital allocated the 29 seed members to various nursing stations to not only guide end users on system operation but also provide immediate assistance to remove barriers when possible. Meanwhile, the case hospital had set up a LINE group chat for communication and interaction. The headquarter IT staff, nurse informaticists, senior management of the case hospital, and all users were invited to join the LINE group chat. All users could voice problems and respond to problems in real time through the LINE group. The attention and support of senior management, appointment of the nurse supervisor as the change agent, establishment of a pioneer team, and effective assistance by seed members who were familiar with the system functions and operations to promote the system enabled users to clearly experience the benefits of the system. These measures effectively improved user satisfaction with the new system, and this outcome corroborated the results from the aforementioned data analysis. For the marginal significance of the path of service quality to user satisfaction, the following reasons are proposed. The new system was mainly developed by the headquarter hospital and was appropriately modified and sequentially introduced in response to the needs of individual hospital units. The IT department in the case hospital had only hardware engineers and no software engineers. All system requirements were collected separately by the chief supervisor, who sent the collected data to the headquarter hospital for modifications. After the system had launched, the chief supervisor was also responsible for collecting opinions and providing feedback to the IT department of the headquarter hospital for further modification. The inability of the IT department to immediately and effectively process users' opinions and the poor timeliness were the possible reasons why users were unable to form positive perceptions for the immediacy and sufficiency of the assistance and support provided by the IT department. Furthermore, the service quality dimension primarily explored the immediacy, adequacy, and appropriateness of the service, support, and assistance provided by the IT department to the users. This included the completeness of software and hardware resources. However, in the initial phase of system launch, the end users continued to use old computers, which had poor hardware performance. Together with insufficient wireless hotspots, the connection and usage of the system were restricted. We believe that the above reasons may further explain the poor perception of service quality by end users and the marginal significance between service quality and satisfaction.

Furthermore, for information quality and satisfaction not reaching the significance level, we noted and offered the following explanations. Prior to the launch of the new system, the nursing staff performed their tasks at the bedside while ideally manually recording the medical materials consumed. The records were then sent to the clerk, who made each entry individually in the old inpatient charging system for approval. The charging process was highly task-dependent. The principle of the new system design was to provide support for nurses to perform professional work and reduce their workload. The charging work was divided into two parts. The first part mainly involves clicking the consumption items, and the second part mainly involves the automatic system that functions as the mechanism through which the amount of money being paid is shown for subsequent verification and approval. After the new system was launched, the nurses only had to click on the consumed medical materials on the interface to complete the frontend work of charging. Frontline nurses were only able to see the different types of medical materials shown on the interface. On the interface, the amount of medical materials consumed was not shown to the nurses, and thus, they did not need to worry about the calculation of the chargeable amount and the subsequent approval work. Thus, they could not form a perception about the actual calculation of charges. Further processing of the charges and the information output were only managed by clerks. The information quality dimension was mainly used to determine whether the format, immediacy, and accuracy of the processed information presented by the system could satisfy or meet users' needs or whether the users noticed the changes in these aspects in the new system. In fact, ward clerks were the ones who could most directly perceive the changes in information quality before and after the launch of the new charging system for inpatient medical materials. However, the participants in this study were limited to frontline nurses, and ward clerks were not included. We believe that the above reasons may explain the lack of perception about information quality by nurses, which led to the failure of attaining significance for this path.

A deeper understanding of the factors affecting user satisfaction with the IS will have both research and practical values to improve the effectiveness of IS implementation and management performance. Our study utilized the DeLone and McLean 2003 IS Success Model. "Subjective norm" was chosen as a variable, and "work performance" was chosen as a dependent variable. This was done to understand the factors impacting the ward nurses' satisfaction with the hospital charging system and their work performance. The model of this study was simple, and the overall explanatory power was 68.9%. This also indicated the feasibility of using the IS Success Model to analyze the effectiveness of a medical-related IS. Additionally, this theory could be widely applied to other areas of studies and so provide a concrete picture of other related research. Furthermore, the current status of system implementation was evaluated through the use of research tools, which allowed the staff in managerial positions to have a better understanding of the main reasons why the lack of significant perceptions about service quality involving the IT department was mainly due to poor hardware performance, insufficient wireless hotspots, and inability to resolve user needs or problems in a timely manner. Therefore, according to the results, the initial steps in the subsequent phase of system promotion are to increase the budget for IT, replace old equipment, and expand wireless hotspots to solve the current urgent problems. Thereafter, users should be provided with better service quality, and the accessibility and availability of the system should be improved. This will increase user satisfaction with the system and thus enhance work support for better work performance. With these factors in mind, the data

presented in this study contribute to both practical and research settings.

Limitations

To explore nursing staff members' evaluation of the new system and to verify the work performance, this study applied the IS Success Model as a theoretical foundation. However, as with all studies of this magnitude, there were several research limitations, although this case study revealed useful and reliable findings. First, other potential factors were not included. Furthermore, this study was conducted in a case hospital with ward nurses as research participants. The selection of the participants might limit the generalization of the results of this research. Moreover, the case hospital implemented the charging system to replace the existing manual charging procedure performed by nurses. By informatizing the charging system, the charging procedures could be simplified so that nurses are able to record the treatments accurately and perform automatic charging. The conclusions to be drawn are that enhancing system quality, information quality, and service quality could lead to improvements in users' satisfaction and work performance.

Recommendations

The introduction of the charging system will require not only user-friendly software but also sufficient IT equipment and a stable wireless connection to be established. In this way, the benefits introduced by using the charging system will be maximized.

Conclusions

Nursing care is known for its heavy workload. Implementation of the new IT charging system can alleviate nurses' heavy workload and improve their work efficiency. However, introducing a new system can be a large change to the entire hospital. It may help to break the old habit of using a paper-based approach, which is most familiar to senior hospital staff and administrators. Senior staff and administrators may be strongly suggested to learn how to use various computer-based operations. In other words, implementing the new IT charging system needs the support and encouragement of hospital directors and staff in high managerial positions. The seed members of the new IT charging system will need to pioneer this concept and share their positive experiences about using the new charging system with other hospital units and associates. Thus, smooth and successful promotion of the new IT charging system will be achieved. When the proposed computerized charging system is successfully operated, other medical organizations will be inspired to adopt the same system soon. Thereafter, a nation-wide cloud-based health recording system could possibly be established and operated in the near future. The continuous replication and spread of positive experiences as such encourage and pave the way for the further promotion of the new system in the future. The successful experiences of the case hospital could serve as a reference for other hospitals in developing countries like Taiwan to promote medical and nursing informatization.

Acknowledgments

This study compiled part of the results from an in-house project in the case hospital (NTUHYL107.C023). We are grateful to the case hospital for providing subsidies, questionnaire printing, and assistance of their nurses. In addition, we would like to thank the superintendent of the National Taiwan University Hospital, the nursing informatics specialist, and the staff of the Management Information System Department for interdisciplinary co-operation. The author Min-Chi Liao is a PhD Program Student in Graduate Institute of Industrial Engineering and Management, National Yunlin University of Science and Technology, Taiwan. Also, she has the role of Nurse Supervisor in the Department of Nursing, National Taiwan University Hospital Yunlin Branch, Taiwan. Finally, we would like to thank Dr. Pingyu Liu, who is now an assistant professor at the National Yunlin University of Science and Technology Library, for his helpful comments and suggestions on earlier versions of this paper.

Authors' Contributions

MCL was responsible for the research design, institutional review board (IRB) application, questionnaire design, data collection, data coding, data analysis, and article writing. ICL was responsible for the study design, IRB application, research model and hypothesis construction, questionnaire development, data analysis, article writing, editing, revision, and responses to reviewers.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Factor loading of each dimension.
[DOCX File , 39 KB - mhealth_v8i3e16381_app1.docx]

References

1. Pang TW. Introduction to Medical Informatics. Taiwan: Wagner Publishing; 2010.
2. Lee PC. 2016-2017 National Health Insurance Annual Report. Taiwan: National Health Insurance Administration. Ministry of Health and Welfare; 2016 Dec 01. URL: https://www.nhi.gov.tw/Nhi_E-LibraryPubWeb/Periodical/Periodical.aspx?TML1_ID=4&Comm_Category_SN=0 [accessed 2020-03-10]

3. Chen WC. A study on information education and digital divide in Taiwan. Journal of Cyber Culture and Information Society 2007;13:193-228 [FREE Full text]
4. Glaser J. The electronic health record: a digital divide? Healthc Financ Manage 2007 Oct;61(10):38-40, 42. [Medline: 17953180]
5. Executive Yuan. Investigation report of digital divide in 2008. URL: https://ws.ndc.gov.tw/001/administrator/10/relfile/0/1000/92313455871.pdf [accessed 2009-01-25]
6. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
7. Lin CH, Lee TT. Promoting Nursing Competitiveness: Introduction to the Digital Divide. The Journal of Nursing 2010;57(1):95-99. [doi: 10.6224/jn.57.1.95]
8. Li MF, Hsiao YS, Lu JW, Lin MH. Reducing Omission Rate For Charging Medical Materials: A Case Study Of A Regional Teaching Hospital. Show Chwan Medical Journal 2014;13(1,2):37-46. [doi: 10.3966/156104972014061301007]
9. Tsai SY, Lei EF, Huang HJ, Chen HL. Improving the Inventory Shortage of Surgical Operational Material through Establishment of an Informational System. Leadership Nursing 2015;16(4):76-89.
10. Frandsen A. The role of disciplining/translating accounting practices in patient‐centred care. Intl Jnl Public Sec Management 2010 Jun;23(4):381-391 [FREE Full text] [doi: 10.1108/09513551011047279]
11. Chien CP, Chien TK, Lai WL. Constructing an Internal Control Index System on the Hospital Financial Dimension. Journal of Global Business Operation and Management 2012;4:161-172. [doi: 10.29967/JGBOM.201209.0013]
12. Hou JY, Chang WH, Hsiao CW, Chiang YC. The Application of Power Tracer BIKM to Create a Visual Accounting Chart for Hospital Management-A Case Study of a Regional Hospital in Southern Taiwan. Journal of Healthcare Management 2010;11(3):75-87. [doi: 10.6174/JHM2010.11(3).75]
13. Maskell BH, Kennedy FA. Why do we need lean accounting and how does it work? J Corp Acct Fin 2007 Mar;18(3):59-73 [FREE Full text] [doi: 10.1002/jcaf.20293]
14. Lim FP. Impact of Information Technology on Accounting Systems. AJMAHS 2013 Dec 31;3(2):93-106. [doi: 10.14257/ajmscahs.2013.12.02]
15. Liu CH, Lin IC, Lu JJ, Cai D. A Smartphone App for Improving Clinical Photography in Emergency Departments: Comparative Study. JMIR Mhealth Uhealth 2019 Jul 31;7(7):e14531 [FREE Full text] [doi: 10.2196/14531] [Medline: 31368444]
16. Chao YL, Lee TT. A Study of the Satisfaction of Nurses with the Clinical Information System in Intensive Care Units. Journal of Nursing and Healthcare Research 2015;11(2):109-118. [doi: 10.6225/JNHR.11.2.109]
17. Huang HY, Chou SS, Tseng KJ, Feng RC. Experiences of Bar Code Medication Administration Deployment and Total Quality Management Strategy. Journal of Healthcare Management 2011;12(2):130-142. [doi: 10.6174/JHM2011.12(2).130]
18. Grabski SV, Leech SA, Schmidt PJ. A Review of ERP Research: A Future Agenda for Accounting Information Systems. Journal of Information Systems 2011 Mar;25(1):37-78 [FREE Full text] [doi: 10.2308/jis.2011.25.1.37]
19. Ives B, Olson M, Baroudi J. The measurement of user information satisfaction. Commun ACM 1983;26(10):785-793 [FREE Full text] [doi: 10.1145/358413.358430]
20. Seddon PB. A Respecification and Extension of the DeLone and McLean Model of IS Success. Information Systems Research 1997 Sep;8(3):240-253 [FREE Full text] [doi: 10.1287/isre.8.3.240]
21. Delone WH, McLean ER. The Delone and McLean Model of Information Systems Success: A Ten-Year Update. Journal of Management Information Systems 2003;19(4):9-30 [FREE Full text]
22. DeLone WH, McLean E. Information Systems Success: The Quest for the Dependent Variable. Information Systems Research 1992 Mar;3(1):60-95 [FREE Full text] [doi: 10.1287/isre.3.1.60]
23. Seddon PB, Kiew MY. A Partial Test and Development of Delone and Mclean's Model of IS Success. AJIS 1996 Nov 01;4(1):90-109 [FREE Full text] [doi: 10.3127/ajis.v4i1.379]
24. Pitt LF, Watson RT, Kavan CB. Service Quality: A Measure of Information Systems Effectiveness. MIS Quarterly 1995 Jun;19(2):173. [doi: 10.2307/249687]
25. Nelson RR, Todd PA, Wixom BH. Antecedents of Information and System Quality: An Empirical Examination Within the Context of Data Warehousing. Journal of Management Information Systems 2014 Dec 08;21(4):199-235 [FREE Full text] [doi: 10.1080/07421222.2005.11045823]
26. Rainer RK, Watson HJ. The Keys to Executive Information System Success. Journal of Management Information Systems 2015 Dec 11;12(2):83-98 [FREE Full text] [doi: 10.1080/07421222.1995.11518082]
27. Parasuraman A, Zeithaml VA, Berry LL. SERVQUAL: A Multiple-item Scale for Measuring Consumer Perceptions of Service Quality. Journal of Retailing 1988;64(1):12-44.
28. Huang MY, Chen TL, Chen YW. Study of Tourism Website Consumers-The Application of Information System Success Model. East-Asia Review 2014(486):31-51.
29. Hwang HG, Chen RF, Lai YC. Factors Affecting User Satisfaction with Data Warehouse Applications of Taiwanese Banking Industry. Journal of E-Business 2010;12(4):693-716.
30. Urbach N, Mueller B. The updated DeLone and McLean Model of information systems success. Information System Theory 2011;1:1-18. [doi: 10.1007/978-1-4419-6108-2_1]

31. Lau F, Price M, Boyd J, Partridge C, Bell H, Raworth R. Impact of electronic medical record on physician practice in office settings: a systematic review. BMC Med Inform Decis Mak 2012 Feb 24;12:10 [FREE Full text] [doi: 10.1186/1472-6947-12-10] [Medline: 22364529]
32. Bossen C, Jensen LG, Udsen FW. Evaluation of a comprehensive EHR based on the DeLone and McLean model for IS success: approach, results, and success factors. Int J Med Inform 2013 Oct;82(10):940-953. [doi: 10.1016/j.ijmedinf.2013.05.010] [Medline: 23827768]
33. Hsieh PJ, Su YH. A Study Exploring Critical Success Factors in the Electronic Medical Record Information System for Medical Information Staff: the Extension of a Successful Model System. Journal of Medical and Health Information Management 2015;13(2):19-39.
34. Chang HF, Lin IC. Modeling the impact of nursing shifting system on job performance. The Journal of Taiwan Association for Medical Informatics 2019;28(1):9-22.
35. Huang WM, Chen LF, Su YH, Chen CC. The influence on shift-to-shift report information system for the performance and satisfaction. 2016 Apr 23 Presented at: The 10th International Conference on Advanced Information Technologies-The 6th Forum in Taiwan Association for Web Intelligence Consortium; April 23, 2016; Taiwan p. 949-968 URL: https://mi.oit.edu.tw/m/404-1027-19009.php?Lang=zh-tw
36. Petter S, Fruhling A. Evaluating the success of an emergency response medical information system. Int J Med Inform 2011 Jul;80(7):480-489. [doi: 10.1016/j.ijmedinf.2011.03.010] [Medline: 21501969]
37. Garcia-Smith D, Effken JA. Development and initial evaluation of the Clinical Information Systems Success Model (CISSM). Int J Med Inform 2013 Jun;82(6):539-552. [doi: 10.1016/j.ijmedinf.2013.01.011] [Medline: 23497819]
38. Campbell JP, McCloy RA, Oppler SH, Sager CE. A theory of performance. In: Personnel Selection. San Francisco: Jossey-Bass; 1993:35-70.
39. Kast F, Rosenzweig J. Organization and Management. In: A Systems and Contingency Analysis. New York: McGraw-Hill; 1985.
40. Hill RJ, Fishbein M, Ajzen I. Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Contemporary Sociology 1977 Mar;6(2):244. [doi: 10.2307/2065853]
41. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211 [FREE Full text] [doi: 10.1016/0749-5978(91)90020-t]
42. Thompson RL, Higgins CA, Howell JM. Influence of Experience on Personal Computer Utilization: Testing a Conceptual Model. Journal of Management Information Systems 2015 Dec 14;11(1):167-187 [FREE Full text] [doi: 10.1080/07421222.1994.11518035]
43. Venkatesh V, Davis FD. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science 2000 Feb;46(2):186-204 [FREE Full text] [doi: 10.1287/mnsc.46.2.186.11926]
44. Venkatesh V, Morris MG. Why Don't Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior. MIS Quarterly 2000 Mar;24(1):115-139 [FREE Full text] [doi: 10.2307/3250981]
45. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425-478 [FREE Full text] [doi: 10.2307/30036540]
46. Garcia-Smith D, Effken JA. Development and initial evaluation of the Clinical Information Systems Success Model (CISSM). Int J Med Inform 2013 Jun;82(6):539-552. [doi: 10.1016/j.ijmedinf.2013.01.011] [Medline: 23497819]
47. Kuo CY, Zheng SP, Zhang ST, Lai XL, Chan CL. Applying the Conceptual Framework of ?Combined Technology Acceptance Model and Theory of Planned Behavior (C-TAM-TPB) ?to Implement Electronic Medical Records: An Example in a Regional Teaching Hospital. Journal of Medicine and Health 2018;7(1):65-75.
48. Huang WM, Wang YH, Hsu HS. An Empirical Study on the Critical Factors in Public Health Information Systems-Taking the Integrated Delivery System with Diabetes Mellitus (IDS-DM) in Central Taiwan for Example. Journal of Innovation and Management 2012;9(3):105-141. [doi: 10.30118/JIM.201208.0005]
49. Chou SS, Chen YJ, Shen YT, Yen HF, Kuo SC. Implementation and Effectiveness of a Bar Code-Based Transfusion Management System for Transfusion Safety in a Tertiary Hospital: Retrospective Quality Improvement Study. JMIR Med Inform 2019 Aug 26;7(3):e14192 [FREE Full text] [doi: 10.2196/14192] [Medline: 31452517]
50. Sowan AK, Leibas M, Tarriela A, Reed C. Nurses' Perceptions of a Care Plan Information Technology Solution With Hundreds of Clinical Practice Guidelines in Adult Intensive Care Units: Survey Study. JMIR Hum Factors 2019 Feb 12;6(1):e11846 [FREE Full text] [doi: 10.2196/11846] [Medline: 30747713]
51. van Dulmen S, Driesenaar JA, van Weert JC, van Osch M, Noordman J. PatientVOICE: Development of a Preparatory, Pre-Chemotherapy Online Communication Tool for Older Patients With Cancer. JMIR Res Protoc 2017 May 10;6(5):e85 [FREE Full text] [doi: 10.2196/resprot.6979] [Medline: 28490421]
52. Ministry of Health and Welfare. Statistics of Nursing Staff in Taiwan December 2019. Taiwan: Taiwan Union of Nurses Association (TUNA); 2019 Dec 2. URL: http://www.nurse.org.tw/publicUI/H/H102.aspx [accessed 2020-03-10]

Abbreviations

EHR: electronic health record
EMR: electronic medical record
HIS: hospital information system
IT: information technology
IS: information system
NHI: National Health Insurance

Edited by G Eysenbach; submitted 26.09.19; peer-reviewed by MB Soomro, A Parush; comments to author 24.11.19; revised version received 25.01.20; accepted 22.02.20; published 25.03.20.

Please cite as:
Liao MC, Lin IC
Performance Evaluation of an Information Technology Intervention Regarding Charging for Inpatient Medical Materials at a Regional Teaching Hospital in Taiwan: Empirical Study
JMIR Mhealth Uhealth 2020;8(3):e16381
URL: http://mhealth.jmir.org/2020/3/e16381/
doi:10.2196/16381
PMID:32209534

Original Paper

Patient Attitudes Toward Mobile Device Use by Health Care Providers in the Emergency Department: Cross-Sectional Survey

Mohamad Alameddine[1], PhD; Hani Tamim[2], PhD; Dima Hadid[3], MPH; Mohamad-Ali Cheaito[3], MD; Maha Makki[2], MPH; Hadi Maatouk[3], MSc; Eveline Hitti[3], MD, MBA

[1]Department of Health Management and Policy, American University of Beirut, Beirut, Lebanon

[2]Department of Internal Medicine, American University of Beirut, Beirut, Lebanon

[3]Department of Emergency Medicine, American University of Beirut, Beirut, Lebanon

Corresponding Author:
Eveline Hitti, MD, MBA
Department of Emergency Medicine
American University of Beirut
Riad El Solh 1107 2020
Beirut,
Lebanon
Phone: 961 1350000
Email: hittieveline@gmail.com

Abstract

Background: Health care provider usage of mobile devices is increasing globally; however, there is little understanding of patient perceptions on this behavior in a health care setting.

Objective: The aim of this study was to assess patients' attitudes toward mobile device usage by health care providers in the emergency department and to identify predictors of these attitudes.

Methods: The study was carried out at the emergency department of a large academic tertiary care medical center in Lebanon. A cross-sectional survey design was adopted by administering a questionnaire to medically stable adult patients who presented to the emergency department with an emergency severity index of 3, 4, or 5 between January 2017 and March 2018. The questionnaire collected relevant patient demographic information and included questions related to their mobile device usage along with those evaluating attitudes for the use of mobile devices by health care providers with respect to six major domains: role in health care, distraction potential, impact on communication, empathy, privacy, and professionalism. The attitude toward mobile device usage by health care providers in the emergency department was the main outcome variable. A stepwise logistic regression model was used to assess the association between the outcome variable and the demographic and attitude-related independent variables.

Results: Among the 438 eligible patients, 338 patients responded to the questionnaire for a response rate of 70.0%. Overall, 313/338 (92.6%) respondents agreed that mobile devices improve health care delivery, whereas 132/338 (39.1%) respondents were opposed to their usage by health care providers in the emergency department (95% CI: 34.0-44.4). The majority (240/338, 71.0%) of patients agreed that mobile devices are a source of distraction to health care providers in the workplace. Females (odds ratio [OR]=1.67, 95% CI: 1.00-2.78) as well as all patients (OR=2.54, 95% CI 1.36-4.76) who believed that mobile devices were a source of distraction, reflecting a lack of professionalism (OR=2.77, 95% CI 1.59-4.82) and impacting the provider's ability to relate to the patient (OR=2.93, 95% CI 1.72-4.99), were more likely to agree that mobile devices should not be used in the emergency department.

Conclusions: Patients' negative attitude toward mobile device use in the emergency department is largely driven by patient gender (females), patient perception of the distraction potential of the devices, and their negative impact on the health care provider's empathy and professionalism. The findings of this study shed light on the importance of encouraging stakeholders to impose a digital professionalism code of conduct for providers working in acute health care settings.

(JMIR Mhealth Uhealth 2020;8(3):e16917) doi:10.2196/16917

KEYWORDS

smart devices; emergency department; patients; attitudes; digital professionalism; code of conduct; empathy; professionalism; distraction; attention

Introduction

The penetration and usage of mobile devices is increasing globally, with growing penetration into the health care sector. A previous survey indicated that 87% of health care providers report using some form of portable network-enabled electronic device such as smartphones in the workplace, paralleling the rapid growth in health care apps, which is now the third fastest–growing app category on the market [1]. Mobile devices are generally considered to be of value to patients and providers in the form of speed of information transmission, clinical decision making, and accessibility [1,2]; however, little is known about patient perceptions of health care providers' usage of mobile devices in a health care setting.

Physicians, both senior and those in training, regularly use mobile devices to access medical apps (eg, drug guides, medical calculators), capture work-related images, respond to bleeps, communicate with their teams, and request diagnostics, among other uses [2-5]. In addition, mobile devices increase provider accessibility, improve communication, and help promote collegiality among the health care delivery team [6]. The usage of mobile devices in health care settings is not a passing trend but rather a practice that is now highly integrated into the work culture, which is likely to expand and grow in the future [7].

Despite these advantages, there are concerns about the regular use of mobile devices in health care settings, including the potential of jeopardizing the privacy of patient information, interfering with patient devices (eg, electromagnetic resonance), as well as cross-contaminating patient care areas [8,9]. Furthermore, there is a potential impact of such distraction on clinical care, especially in high-risk areas such as the emergency department characterized by a high cognitive load and regular interruptions. Such distraction potential has serious safety implications, including the risk that physicians and residents will miss vital patient information [8,10]. In fact, approximately 41% of health care providers at the American University of Beirut Medical Center (Beirut, Lebanon) reported distraction by nonwork-related use of mobile devices [11]. There is a growing body of literature on the safety implications of the usage of mobile devices and the so-called "inattentional blindness" associated with their use. However, few studies have directly investigated patient perceptions of their providers' usage of mobile devices in health care settings or the impact of such use on the physician-patient relationship.

Studies that have examined the effect of mobile device usage on interpersonal relations showed that the presence of mobile devices in a social setting negatively affects the quality of conversations, extent of satisfaction with a social encounter, as well as the level of empathy and connection [12-14]. Several studies have explored patient attitudes toward health technology and their impact on the patient-provider relationship. However, these studies have been limited to the use of computers in the consultation room [15-17], tablets in the examination room and telerounding [18,19], mobile phone images in wound care [20-22], personal digital assistants in emergency departments [23], and personalized smart bedside stations in an inpatient setting [24]. In general, most patients did not express a negative attitude toward their physicians' use of such technology [18,23,25] and did not feel that their interaction with their care provider was less personal due to the use of the technology [16,18].

However, these findings cannot be extrapolated to the usage of mobile devices for several reasons, including the mobility of such devices, their strong distracting potential, association with the users' wider social network even when not actively being used, and accumulating evidence on their negative impact on empathy and quality of interactions. Focused assessments of patient attitudes toward their health care providers' use of mobile devices in health care settings is of importance [12-14], especially in the context of the emergency department. This is because the nature of work in the emergency department poses specific challenges for both safety and relationship building, which is characterized by high volumes, short interactions, and frequent interruptions [26-28]. Therefore, understanding patient perspectives in this setting can guide policy and practice recommendations that will help address patient concerns and preserve the patient-provider relationship as mobile device adoption in health care continues to rise.

Accordingly, the aim of this study was to assess patients' attitudes toward the usage of mobile devices by health care providers in the emergency department. Moreover, we statistically explored the predictors of these attitudes, including demographic characteristics as well as perceptions of the role in health care, distraction potential, impact on communication, empathy, privacy, and professionalism.

Methods

Study Design and Setting

The study was conducted at the emergency department of an academic tertiary-care medical center in Lebanon (American University of Beirut Medical Center, Beirut, Lebanon). The annual census of 2016 showed that the emergency department received 55,000 patient visits, with 70% of patients presenting on weekdays and the remaining presenting on weekends. Trained nurses triage patients based on the emergency severity index (ESI) and age. ESI is a 5-level index used in emergency departments to rate a patient's acuity from level 1 (most urgent) to level 5 (least urgent) based on an estimation of resources required [29].

Desktop computers are available at the nursing station for electronic-based ordering of labs and diagnostics. Otherwise, all other documentation, including nursing and medication orders, are paper-based. There are no workstations on wheels and no tablets integrated into any of the documentation or ordering workflows. All of the health care providers in this

emergency department report bringing a mobile device to work, with 83 out of 97 respondents (86%) reporting use of a mobile device for medical purposes [11].

A cross-sectional survey was administered to adult patients presenting to the emergency department. Data collection was carried out between January 2017 and March 2018. Ethical approval was obtained from the Institutional Review Board of the American University of Beirut that the medical center is affiliated with (protocol number ED.EH.06).

Recruitment of Participants

The eligibility criteria were patients aged 18 years and older presenting to the emergency department during the study period with an ESI of 3, 4, or 5; present in the emergency department for a minimum of 2 hours; and the primary attending physician determined that they were medically stable. Among the 483 eligible subjects invited to participate in this study, 145 refused to participate, resulting in a total of 338 subjects available for analyses (response rate of 70.0%). The main reasons for refusal to participate included not feeling well enough to participate, lack of interest in the research study, or not having time to participate.

A stratified random sampling design was adopted, in which the strata were defined as weekdays (237/338, 70.1%) vs weekends (101/338, 29.9%). The random aspect of this sampling was achieved by carrying out data collection during different times and days of the week, in which all eligible patients available during the visit were identified using the emergency department dashboard, which is an in-house electronic patient tracking system that includes the patients' age, gender, ESI, and arrival time.

Measurements

A mobile device was defined as any handheld portable network-enabled electronic device that is generally connected to other devices or networks via different wireless protocols [30-32], primarily smartphones, which have not been addressed by other studies.

A survey instrument in English was developed to evaluate patients' awareness and attitudes to the utilization of mobile devices in the emergency department. A review of the published peer-reviewed literature and other surveys examining the use mobile devices was carried out to develop the survey questionnaire used in this study [10]. Upon this review, a preliminary version of the survey was constructed, which was further customized to the institutional setting and reviewed by a group of experts, including a statistician, the director of the Emergency Medicine Department, a health management and policy expert, and a social scientist, to enhance content validity (see Multimedia Appendix 1). The English version of the questionnaire was translated into Arabic and then back-translated to English, and the two drafts were compared for consistency. The preliminary drafts were then pilot-tested among 45 patients (who were excluded from the final analyses) for redundancy, validity, and clarity of the questions and statements. The survey was subsequently revised and modified based on patient feedback. Patients were given the option to choose which version they would like to complete based on their preferences. Before administering the survey, the participants were asked to read and sign an informed consent form.

The survey included relevant patient demographic information (age, gender, level of education, patient arrival time, employment status, and monthly income) and their usage of mobile devices. The questionnaire also included a list of statements graded on a 4-point Likert scale ("disagree," "strongly disagree," "strongly agree," or "agree") that evaluated patients' attitude toward the use of mobile devices by health care providers with respect to six major domains: role in health care, distraction potential, impact on communication, empathy, privacy, and professionalism.

The attitude toward the usage of mobile device in the emergency department was the main outcome variable considered in this study. More specifically, the statement was "I believe mobile devices should not be used by health care providers in emergency departments." Responses were divided into two groups: agree (those who answered "agree" or "strongly agree" to that question) and disagree (those who answered "disagree" or "strongly disagree" to that question).

Sample Size Calculation

Sample size calculation was carried out considering the primary outcome and based on a previous study carried out in a pediatric teaching hospital in Australia, which reported that 33% of patients were against the use of mobile devices at bedside [33]. A sample of 338 patients was estimated with a 95% CI and 5% margin of error to detect a similar distribution.

Statistical Analysis

Statistical Package for Social Sciences version 24.0 (SPSS Inc, Chicago, IL, USA) was used for data cleaning, management, and analyses. Descriptive statistics are summarized by the number and percentage for categorical variables. The association between "mobile devices should not be used in the emergency department" and other categorical variables was assessed using the Chi square test. Multivariate regression analysis was performed to adjust for potentially confounding variables. Stepwise logistic regression analysis was used to assess the association between the response to "mobile devices should not be used in the emergency department" as a binary variable (agree versus disagree) with all demographic variables and the statistically significant attitude variables. $P<.05$ was set as the entry threshold of potential predictors into the model, whereas $P<.10$ was set as the threshold for removal from the model. The results are presented as the odds ratio (OR) and 95% CI; $P<.05$ was considered statistically significant.

Results

Among the 338 respondents, 132 (39.1%) were opposed to the usage of mobile devices by health care providers in the emergency department. Table 1 presents the demographic characteristics of all patients and the self-reported mobile device usage, as well as the association with the main outcome (health care providers should not use mobile devices in the emergency department). Overall, the study sample was relatively young with 174/338 (51.5%) respondents aged 35 years or less, with a slightly higher number of women. The majority of the patients

were employed and had completed at least a university degree, with slightly more than half earning more than 2000 USD per month. The large majority of respondents reported owning a mobile device, most commonly a smartphone, with the top uses including messaging apps, phone calls, and social media. Analysis of the association between the main outcome and different variables (demographic and self-reported mobile device usage) revealed gender as the only significant factor, with females more likely to agree that mobile devices should not be used in the emergency department.

Table 1. Demographic characteristics, self-reported usage of a mobile device, and their association with main outcome.

Characteristic	Health care providers should not use a mobile device in the emergency department, n (%)			
	All (N=338)	Disagree (n=206)	Agree (n=132)	*P* value
Demographic				
Gender				.02
Male	158 (46.7)	107 (51.9)	51 (38.6)	
Female	180 (53.3)	99 (48.1)	81 (61.4)	
Age (years)				.53
<25	81 (24.0)	48 (23.3)	33 (25.0)	
25-35	93 (27.5)	57 (27.7)	36 (27.3)	
36-50	73 (21.6)	45 (21.8)	28 (21.2)	
51-65	52 (15.4)	28 (13.6)	24 (18.2)	
66+	39 (11.5)	28 (13.6)	11 (8.3)	
Education level				.78
Less than high school	32 (9.6)	20 (10.0)	12 (9.2)	
High school graduate	47 (14.2)	26 (12.9)	21 (16.0)	
University graduate	194 (58.4)	121 (60.2)	73 (55.7)	
Postgraduate	59 (17.8)	34 (16.9)	25 (19.1)	
Employed	221 (65.4)	133 (64.9)	88 (66.7)	.74
Monthly combined family income (USD)				.40
<1000	42 (19.9)	30 (22.7)	12 (15.2)	
1000-2000	55 (26.1)	34 (25.8)	21 (26.6)	
2000+	114 (54.0)	68 (51.5)	46 (58.2)	
Utilization				
Own a mobile device	327 (96.7)	200 (97.1)	127 (96.2)	.76
Type of mobile device owned				
Smartphone	319 (97.6)	195 (97.5)	124 (97.6)	>.99
Tablet	109 (33.3)	70 (35.0)	39 (30.7)	.42
Smartwatch/band	27 (8.3)	21 (10.5)	6 (4.7)	.06
Regular phone	1 (0.3)	1 (0.5)	0 (0.0)	>.99
Other	95 (29.1)	57 (28.5)	38 (29.9)	.81
Reasons for using a mobile device				
Phone calls	284 (86.9)	176 (88.0)	108 (85.0)	.30
Messaging apps	284 (86.9)	175 (87.5)	109 (85.8)	.66
Social media	205 (62.7)	124 (62.0)	81 (63.8)	.75
Games	112 (34.3)	76 (38.0)	36 (28.3)	.07
Browsing the internet	202 (61.8)	120 (60.0)	82 (64.6)	.41

Table 2 presents the descriptive analyses of patients' attitudes toward the usage of a mobile device by health care providers in the emergency department, as well as the association with the main outcome. The majority of respondents believed that mobile devices play a role in patient care and improve health care delivery, but that they should only be used for medical care

purposes. According to the respondents, the top reasons for appropriate use of mobile devices in a health care setting are accessing medical information, sending/receiving medical documents/images, looking up patient information, and communicating via messaging apps.

In addition, two thirds of respondents reported that the use of mobile devices does not demonstrate a lack of professionalism, and more than half believe that the use of mobile devices does not cause a breach of confidential patient information. By contrast, more than two thirds of respondents agreed that mobile devices are a distraction to health care providers in the workplace, half agreed that the use of mobile devices by health care providers leads to poor patient-provider communication, and close to half agreed that the use of mobile devices impacts the ability of health care providers to relate to patients.

Moreover, the large majority of patients who agreed that mobile devices are a distraction to health care providers in the workplace were more likely to agree that mobile devices should not be used. Consistently, most patients who agreed that mobile devices lead to poor patient-provider communication were also more likely to agree that mobile devices should not be used in the emergency department. Moreover, patients who agreed that use of mobile devices impacts providers' ability to relate to patients, demonstrates lack of professionalism, and causes a breach of confidential information were more likely to agree that mobile devices should not be used in the emergency department. Finally, those who do not like providers using their mobile devices when treating them were more likely to agree that mobile devices should not be used in the emergency department.

Table 2. Descriptive analysis of patients' attitudes toward the usage of a mobile device by health care professionals and association with main outcome.

Attitude	Health care providers should not use mobile devices in the emergency department, n (%)			
	All (N=338)	Disagree (n=206)	Agree (n=132)	*P* value
Role in health care				
Agree that mobile devices play a role in patient care	279 (85.3)	176 (88.0)	103 (81.1)	.09
Mobile device functions in hospital setting				
Access medical information (general)	249 (76.1)	155 (77.5)	94 (74.0)	.47
Send/receive medical documents/images	245 (74.9)	155 (77.5)	90 (70.9)	.18
Look up patient information	213 (65.1)	131 (65.5)	82 (64.6)	.86
Personal calls	150 (45.9)	95 (47.5)	55 (43.3)	.46
Messaging apps	153 (46.8)	94 (47.0)	59 (46.5)	.92
Facebook or other social media	105 (32.1)	58 (29.0)	47 (36.2)	.13
Mobile devices play a role in improving health care delivery	313 (92.6)	196 (95.1)	117 (88.6)	.03
Mobile devices should only be used for medical care	296 (87.8)	183 (89.3)	113 (85.6)	.32
Distraction potential				
Mobile devices are a distraction to health care providers	240 (71.0)	126 (61.2)	114 (86.4)	<.001
Health care providers spend more time on their mobile devices than with me	13 (3.8)	6 (2.9)	7 (5.3)	.27
Communication and empathy				
Mobile device usage by health care providers leads to poor patient-provider communication	170 (50.3)	81 (39.3)	89 (67.4)	<.001
Health care providers' mobile devices usage impacts their ability to relate to me	151 (44.8)	61 (29.6)	90 (68.7)	<.001
Professionalism and privacy				
I don't like health care providers using their mobile devices when treating me	205 (60.7)	93 (45.1)	112 (84.8)	<.001
Mobile device usage demonstrates a lack of professionalism	109 (32.3)	39 (18.9)	70 (53.4)	<.001
Mobile device usage causes a breach of confidential information	138 (40.9)	68 (33.0)	70 (53.4)	<.001

Table 3 summarizes the independent factors associated with patients' believing that mobile devices should not be used in the emergency department. Women were more likely to agree that mobile devices should not be used in the emergency department. In addition, patients who agreed that mobile devices were a source of distraction and those who believed they reflected lack of professionalism were more likely to agree that mobile devices should not be used in the emergency department. Moreover, those who felt that the usage of mobile devices impacted the provider's ability to relate to them were more

likely to agree that mobile devices should not be used in the emergency department.

Table 3. Multivariate regression analysis for predictors of the main outcome[a].

Predictor variable	Health care providers should not use mobile devices in the emergency department (reference: Disagree)	
	Odds ratio (95% CI)	*P* value
Gender	1.67 (1.00-2.78)	.05
Distraction to health care provider	2.54 (1.36-4.76)	.03
Demonstrates lack of professionalism	2.77 (1.59-4.82)	<.001
Impacts health care provider's ability to relate to me	2.93 (1.72-4.99)	<.001

[a]The following variables were included in the full model: gender (reference: male); age (reference: <25 years); education (reference: <high school); employed (reference: unemployed); distraction; improves health care delivery; mobile device should be used only for medical care; poor communication; impacts health care providers' ability to relate to me; lack of professionalism; causes a breach of confidential information.

Discussion

Principal Findings

This study represents a rare attempt to examine patient perspectives on the use of mobile devices in an acute health care setting. With the growing body of literature on the distraction potential of mobile devices and their negative impact on interpersonal relationships, understanding patients' perceptions toward the use of mobile devices by health care providers is important for curbing any unintended consequences of their permeation into health care. The present findings reveal that although the majority of patients agree that mobile devices can improve health care and should be used for medical purposes, many felt that mobile devices should not be used in the emergency department (41%).

The results of this study concur with those of existing literature showing that patients acknowledge the importance of technology usage in health care delivery [15-25]. The surveyed patients had clear views on the use of mobile devices, with the majority stating that mobile devices improve health care delivery (92.5%) and that they should be used for medical care (87.5%). Furthermore, close to three quarters of the respondents believed that physicians use their mobile devices in health care settings to access medical information and send or receive medical documents.

The overall positive attitude expressed by patients is counterbalanced by several concerns. Many patients felt that mobile device usage impacted the providers' ability to relate to them. This is in contrast to studies that considered the patient perspective on other forms of technology, including computers, tablets, and personal digital assistants, who denied any change or depersonalization in their interaction with physicians using such devices [15-20,24,25]. Most previous studies reported that the use of technology in health care is considered to enhance communication and quality of care [19,21,25]. However, our findings are in line with the psychology literature suggesting that mobile devices may negatively affect relationships by dividing an individual's attention between an immediate face-to-face interaction and a distant wider social network, even when not being actively used. Studies also show that the mere presence of a mobile device during a paired interaction inhibits the ability to develop closeness and trust, in addition to reducing perceived empathy by the partner [13]. Within the health care context, specific features of mobile devices, compared to other technologies, may accentuate the abovementioned negative feelings for patients. The mobility and accessibility of mobile devices, along with their high distraction and addiction potential [34,35], as well as the reduced visibility of the mobile device screen may all heighten feelings of isolation and suspicions that the mobile device is distracting providers from patient care and the face-to-face interaction [34]. This finding is particularly relevant to the context of emergency departments where rapport building, assessment, and communication about care are all being squeezed into a brief encounter where the provider and patient are often meeting for the first time, in a high-interruption, high-anxiety, and high-risk setting [27,28].

Distraction also emerged as one of the main concerns in our study, with more than two thirds of the respondents reporting that mobile devices may be a source of distraction to health care providers in the workplace (71.1%). This finding is in contrast to studies considering the patient perception of provider usage of computers where distraction did not emerge as a significant patient concern. Our findings instead concur with studies that showed high rates of health care professional self-reporting of distraction by a mobile device [10,15,36]. The link between the use of mobile devices and distractibility has been extensively established, showing an association with reduced reaction time, the worst performance on tasks that require cognitive focus, as well as "inattention blindness", which is the reduced ability to notice unique and novel stimuli [8,37-41]. This perceived distraction may also add to patient concerns about providers relating to them or displaying empathy during the episode of care.

The abovementioned concerns can explain the fact that 2 out of every 5 patients (41%) felt that mobile devices should not be used in the emergency department. Although age, education, and income level were not associated with this patient opinion, gender did emerge as a significant predictor variable, with women being significantly more likely to disagree with the usage of mobile devices in the emergency department setting. This finding is in line with studies showing that men, as compared to women, are more likely to find talking on their mobile phones in various personal situations more acceptable

[42]. Moreover, patient perception of how mobile device usage may influence factors related to the care process was highly associated with the potential impact on the provider-patient relationship. Specifically, those who believed that the use of mobile devices reflected lack of professionalism, distracted from care, and negatively impacted the provider's ability to relate to them also agreed that mobile devices should not be used in the emergency department. Such findings suggest that providers were visible to patients while using their mobile devices or that such use may have taken place during the clinical encounter with the patient. The use of mobile devices by health care providers during clinical encounters would be perceived by many patients to reflect a lack of professionalism, whether it took place inside the emergency department or any other care setting. However, the physical layout of the emergency department, in which all patient care areas are open to the central health care provider station, may have accentuated this perception since health care providers are continuously visible to patients. The ability to relate to patients, a reflection of empathy, was the strongest driver for disagreeing with usage in the emergency department, corroborating the psychology literature on the impact of using mobile devices on empathy and relationship building [12-14].

This study sheds light on serious patient concerns that warrant consideration as mobile device permeability in health care continues to grow. Multiple sectors have already addressed the distracting potential of mobile devices on safety through regulatory initiatives such as the Distracted Driver Law that prohibits usage while driving and the "Sterile cockpit law" that prohibits pilots and crew members from engaging in any activity during critical phases of takeoff and landing [43]. Although the use of mobile devices has become too intertwined with clinical care for a complete ban to be possible, there is a clear need to place some guidelines surrounding their use in health care and introduce some codes of conduct. From a policy perspective, managers need to ensure the right balance between security and liberty [44]. From a liberty perspective, care providers should have the freedom to use their devices as they deem appropriate. From a security perspective, such use should be regulated to mitigate the negative impact on providers' productivity, patient safety, and the patient-provider relationship. The tradeoff between security and liberty is inevitable and would call for wider discussions between care providers, administrators, regulatory bodies, and the ministry of health to build support for regulatory policies and procedures that could be endorsed at the national level with the support of concerned stakeholders.

Gill et al [8] proposed several guidelines that aimed at securing institutional networks and regulating the use of mobile devices for nonwork-related purposes. This can be realized through ensuring high-security wifi connections and extending firewalls to identify and control the use of an application on the network. Limiting access to social network websites such as YouTube and Facebook and establishing an intracompany communication network are also possible solutions [8].

Nevertheless, successful interventions cannot solely rely on technological solutions to limit personal use. Developing and nurturing a digital professionalism code of conduct will be essential. Setting expectations for clinical care usage in clinical areas and designating separate hotspots for personal use should be part of this implementation. Raising awareness on the impact of the use of mobile devices on face-to-face interactions, empathy, and communication is also essential, with specific attention to female patients. Similarly, it is important to develop best practices for the use of mobile devices around patients, including maintaining eye contact, and explaining to patients why they are using the device to counter the limited screen visibility and associated suspicions that arise.

Although the external validity of the findings and recommendations of this study are particularly applicable to the emergency department setting, they are also applicable, with some contextualization, to other care areas within a health care institution. Future studies should validate the present findings in other care areas and across other contexts.

Limitations

The results of our study should be considered in light of its limitations. First, although the research team assured patients with the confidentiality of their responses and that their responses will not impact the care they are receiving, there remains a risk of a social desirability bias with patients potentially modifying their responses to prevent putting the care providers at risk. Second, the cross-sectional nature of the study is only able to discover associations, and it is difficult to establish causal relationships, leading to a risk of possible spurious associations. In addition, we excluded high-acuity patients (ESI 1 and 2), which may have led to response bias. However, patients with an ESI 3, 4, and 5 collectively comprise 80% of our population. Lastly, the nature of the study as a single-center assessment with a specific patient base may affect the generalizability to other patient populations.

Conclusion

Patients in the emergency department recognize the important role of mobile devices in health care delivery and patient care. Nonetheless, 2 out of every 5 patients believe that mobile devices should not be used in the emergency department. This seems to be driven by gender, with women more likely to disagree with usage in the emergency department, along with patients' perception of how mobile devices may negatively impact the fundamentals of care and the patient-provider relationship, namely professionalism, provider attention, and their ability to relate to patients. This is particularly important in the emergency department setting, where time constraints challenge a physician's ability to build a rapport with patients. Accordingly, this study highlights the significance of fostering and cultivating, in consultation with concerned stakeholders, a digital professionalism code of conduct in the emergency department with particular attention to female patients.

Acknowledgments

This work was supported by the Medical Practice Plan (MPP) grant (MPP 11. 320083.xxxxx.11465.720.9999.0000).

Authors' Contributions

EH conceptualized the study, provided insights into the Introduction and Discussion sections, critically revised the manuscript, and read and approved the final manuscript. MA contributed to conceptualization of the topic and the analysis and the interpretation of data, critically revised the manuscript with a focus on the Discussion and contribution to the literature, and read and approved the final manuscript. DH contributed to interpreting the results and the implementation of the study, provided insights into the Discussion section, and contributed to the writing and editing of the manuscript. MC contributed to the literature review and implementation of the study, provided insights into the Discussion section, and contributed to the writing and editing of the manuscript. HT contributed to the analysis and interpretation of data, critically revised the manuscript, and read and approved the final manuscript. MM contributed to the analysis and interpretation of data, and the reviewing and editing of the manuscript. HM contributed to the implementation of the study, reviewing, and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
English Questionnaire.
[PDF File (Adobe PDF File), 996 KB - mhealth_v8i3e16917_app1.pdf]

References

1. Ventola C. Mobile devices and apps for health care professionals: uses and benefits. P & T 2014 May;39(5):356-364 [FREE Full text] [Medline: 24883008]
2. Franko OI, Tirrell TF. Smartphone app use among medical providers in ACGME training programs. J Med Syst 2012 Oct 4;36(5):3135-3139. [doi: 10.1007/s10916-011-9798-7] [Medline: 22052129]
3. O'Connor P, Byrne D, Butt M, Offiah G, Lydon S, Mc Inerney K, et al. Interns and their smartphones: use for clinical practice. Postgrad Med J 2014 Feb 15;90(1060):75-79. [doi: 10.1136/postgradmedj-2013-131930] [Medline: 24243966]
4. Tsai H, Pong Y, Liang C, Lin P, Hsieh C. Teleconsultation by using the mobile camera phone for remote management of the extremity wound: a pilot study. Ann Plast Surg 2004 Dec;53(6):584-587. [doi: 10.1097/01.sap.0000130703.45332.3c] [Medline: 15602257]
5. Hsieh C, Tsai H, Yin J, Chen C, Yang JC, Jeng S. Teleconsultation with the mobile camera-phone in digital soft-tissue injury: a feasibility study. Plast Reconstr Surg 2004 Dec;114(7):1776-1782. [doi: 10.1097/01.prs.0000142402.07896.21] [Medline: 15577348]
6. American Medical Association. Professionalism in the use of social media URL: https://www.ama-assn.org/delivering-care/ethics/professionalism-use-social-media [accessed 2019-03-19]
7. Mostaghimi A, Crotty BH. Professionalism in the digital age. Ann Intern Med 2011 Apr 19;154(8):560-562. [doi: 10.7326/0003-4819-154-8-201104190-00008] [Medline: 21502653]
8. Gill P, Kamath A, Gill TS. Distraction: an assessment of smartphone usage in health care work settings. Risk Manag Healthc Policy 2012;5:105-114. [doi: 10.2147/RMHP.S34813] [Medline: 22969308]
9. Chisholm CD, Dornfeld AM, Nelson DR, Cordell WH. Work interrupted: a comparison of workplace interruptions in emergency departments and primary care offices. Ann Emerg Med 2001 Aug;38(2):146-151. [doi: 10.1067/mem.2001.115440] [Medline: 11468609]
10. Katz-Sidlow RJ, Ludwig A, Miller S, Sidlow R. Smartphone use during inpatient attending rounds: prevalence, patterns and potential for distraction. J Hosp Med 2012 Oct 28;7(8):595-599. [doi: 10.1002/jhm.1950] [Medline: 22744793]
11. Alameddine M, Soueidan H, Makki M, Tamim H, Hitti E. The Use of Smart Devices by Care Providers in Emergency Departments: Cross-Sectional Survey Design. JMIR Mhealth Uhealth 2019 May 05;7(6):e13614 [FREE Full text] [doi: 10.2196/13614] [Medline: 31199328]
12. Misra S, Cheng L, Genevie J, Yuan M. The iPhone Effect. Environ Behav 2014 Jul;48(2):275-298. [doi: 10.1177/0013916514539755]
13. Przybylski AK, Weinstein N. Can you connect with me now? How the presence of mobile communication technology influences face-to-face conversation quality. J Soc Pers Relat 2012 Jul 19;30(3):237-246. [doi: 10.1177/0265407512453827]
14. Rotondi V, Stanca L, Tomasuolo M. Connecting alone: Smartphone use, quality of social interactions and well-being. J Econ Psychol 2017 Dec;63:17-26. [doi: 10.1016/j.joep.2017.09.001]
15. Rethans J, Höppener P, Wolfs G, Diederiks J. Do personal computers make doctors less personal? Br Med J (Clin Res Ed) 1988 May 21;296(6634):1446-1448 [FREE Full text] [doi: 10.1136/bmj.296.6634.1446] [Medline: 3132287]

16. Aydin C, Rosen P, Jewell S, Felitti V. Computers in the examining room: the patient's perspective. Proc Annu Symp Comput Appl Med Care 1995:824-828 [FREE Full text] [Medline: 8563407]
17. Ridsdale L, Hudd S. Computers in the consultation: the patient's view. Br J Gen Pract 1994 Aug;44(385):367-369 [FREE Full text] [Medline: 8068397]
18. Strayer SM, Semler MW, Kington ML, Tanabe KO. Patient attitudes toward physician use of tablet computers in the exam room. Fam Med 2010 Oct;42(9):643-647 [FREE Full text] [Medline: 20927673]
19. Kaczmarek BF, Trinh Q, Menon M, Rogers CG. Tablet telerounding. Urology 2012 Dec;80(6):1383-1388. [doi: 10.1016/j.urology.2012.06.060] [Medline: 23206790]
20. Sikka N, Carlin KN, Pines J, Pirri M, Strauss R, Rahimi F. The use of mobile phones for acute wound care: attitudes and opinions of emergency department patients. J Health Commun 2012 May 02;17(Supp 1):37-42; quiz 42. [doi: 10.1080/10810730.2011.649161] [Medline: 22548597]
21. Nair AG, Potdar NA, Dadia S, Aulakh S, Ali MJ, Shinde CA. Patient perceptions regarding the use of smart devices for medical photography: results of a patient-based survey. Int Ophthalmol 2019 Apr 6;39(4):783-789. [doi: 10.1007/s10792-018-0878-2] [Medline: 29511959]
22. Kaliyadan F, Amin TT, Kuruvilla J, Ali WHAB. Mobile teledermatology – patient satisfaction, diagnostic and management concordance, and factors affecting patient refusal to participate in Saudi Arabia. J Telemed Telecare 2013 Sep 03;19(6):315-319. [doi: 10.1177/1357633x13501778]
23. Rudkin SE, Langdorf MI, Macias D, Oman JA, Kazzi AA. Personal digital assistants change management more often than paper texts and foster patient confidence. Eur J Emerg Med 2006 Apr;13(2):92-96. [doi: 10.1097/01.mej.0000192049.04729.0c] [Medline: 16525237]
24. Ryu B, Kim S, Lee K, Hwang H, Yoo S. Inpatient satisfaction and usage patterns of personalized smart bedside station system for patient-centered service at a tertiary university hospital. Int J Med Inform 2016 Nov;95:35-42. [doi: 10.1016/j.ijmedinf.2016.09.003] [Medline: 27697230]
25. Houston T, Ray M, Crawford M, Giddens T, Berner E. Patient perceptions of physician use of handheld computers. AMIA Annu Symp Proc 2003:299-303 [FREE Full text] [Medline: 14728182]
26. Roter D. The enduring and evolving nature of the patient–physician relationship. Patient Educ Counsel 2000 Jan;39(1):5-15. [doi: 10.1016/s0738-3991(99)00086-5]
27. Rhodes KV, Vieth T, He T, Miller A, Howes DS, Bailey O, et al. Resuscitating the physician-patient relationship: emergency department communication in an academic medical center. Ann Emerg Med 2004 Sep;44(3):262-267. [doi: 10.1016/j.annemergmed.2004.02.035] [Medline: 15332069]
28. Knopp R, Rosenzweig S, Bernstein E, Totten V. Physician-patient communication in the emergency department, part 1. Acad Emerg Med 1996 Nov;3(11):1065-1069 [FREE Full text] [doi: 10.1111/j.1553-2712.1996.tb03357.x] [Medline: 8922019]
29. Eitel DR. The Emergency Severity Index Triage Algorithm Version 2 Is Reliable and Valid. Acad Emerg Med 2003 Oct 01;10(10):1070-1080. [doi: 10.1197/s1069-6563(03)00350-6]
30. Gillon W, Lei H, Farrand T, Bryan D. Google Patents. 2017. Establishing and managing alternative networks for high quality of service communications URL: https://patents.google.com/patent/US20170034062A1/en [accessed 2019-03-20]
31. D'Ambra J, Wilson C, Akter S. Do e-books enhance the reading experience: An affordance theory perspective. In: Diffusion Interest Group In Information Technology (DIGIT) 2016 Proceedings (pp. 1-9). United States: Association for Information Systems; 2017 Feb 28 Presented at: Twenty-First DIGIT Workshop; December 2016; Dublin, Ireland p. 2016.
32. Selvam R, Shanmugam D, Annadurai S, Rangasamy J. Decomposed S-Boxes and DPA Attacks: A Quantitative Case Study Using PRINCE. 2016 Presented at: International Conference on Security, Privacy, and Applied Cryptography Engineering; December 14-18, 2016; Hyderabad, India. [doi: 10.1007/978-3-319-49445-6_10]
33. Alexander SM, Nerminathan A, Harrison A, Phelps M, Scott KM. Prejudices and perceptions: patient acceptance of mobile technology use in health care. Intern Med J 2015 Nov 13;45(11):1179-1181. [doi: 10.1111/imj.12899] [Medline: 26349003]
34. Jiang J. Pew Research Center. 2018. Millennials stand out for their technology use, but older generations also embrace digital life URL: https://www.pewresearch.org/fact-tank/2019/09/09/us-generations-technology-use/ [accessed 2019-03-20]
35. Hofmann W, Baumeister RF, Förster G, Vohs KD. Everyday temptations: an experience sampling study of desire, conflict, and self-control. J Pers Soc Psychol 2012 Jun;102(6):1318-1335. [doi: 10.1037/a0026545] [Medline: 22149456]
36. Smith T, Darling E, Searles B. 2010 Survey on cell phone use while performing cardiopulmonary bypass. Perfusion 2011 Sep 18;26(5):375-380. [doi: 10.1177/0267659111409969] [Medline: 21593081]
37. Simons DJ, Chabris CF. Gorillas in our midst: sustained inattentional blindness for dynamic events. Perception 1999 Sep 25;28(9):1059-1074. [doi: 10.1068/p281059] [Medline: 10694957]
38. Simons DJ. Attentional capture and inattentional blindness. Trends Cogn Sci 2000 Apr;4(4):147-155. [doi: 10.1016/s1364-6613(00)01455-8]
39. Newby EA, Rock I. Inattentional blindness as a function of proximity to the focus of attention. Perception 1998 Jun 25;27(9):1025-1040. [doi: 10.1068/p271025] [Medline: 10341933]
40. Mack A, Rock I. Inattention blindness. Cambridge, MA: MIT Press; 1998.

41. Hyman IE, Boss SM, Wise BM, McKenzie KE, Caggiano JM. Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. Appl Cognit Psychol 2009 Oct 19;24(5):597-607. [doi: 10.1002/acp.1638]
42. Baron NS, Campbell EM. Gender and mobile phones in cross-national context. Language Sci 2012 Jan;34(1):13-27. [doi: 10.1016/j.langsci.2011.06.018]
43. Chute R. SJSU Scholar Works. 1994 May. Issues in cockpit/cabin communication and coordination. Masters Thesis URL: https://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1733&context=etd_theses [accessed 2020-03-18]
44. Stone DA. Policy Paradox: The Art of Political Decision Making. New York City: WW Norton & Co Inc; 1997.

Abbreviations

ESI: emergency severity index
OR: odds ratio

Edited by G Eysenbach; submitted 06.11.19; peer-reviewed by E Ding; comments to author 28.11.19; revised version received 04.12.19; accepted 06.02.20; published 31.03.20.

Please cite as:
Alameddine M, Tamim H, Hadid D, Cheaito MA, Makki M, Maatouk H, Hitti E
Patient Attitudes Toward Mobile Device Use by Health Care Providers in the Emergency Department: Cross-Sectional Survey
JMIR Mhealth Uhealth 2020;8(3):e16917
URL: http://mhealth.jmir.org/2020/3/e16917/
doi:10.2196/16917
PMID:32229474

Original Paper

Nurse Coaching and Mobile Health Compared With Usual Care to Improve Diabetes Self-Efficacy for Persons With Type 2 Diabetes: Randomized Controlled Trial

Heather M Young[1], PhD; Sheridan Miyamoto[2], PhD; Madan Dharmar[1], PhD; Yajarayma Tang-Feldman[1], MA

[1]Betty Irene Moore School of Nursing, University of California, Davis, Sacramento, CA, United States

[2]College of Nursing, The Pennsylvania State University, University Park, PA, United States

Corresponding Author:
Heather M Young, PhD
Betty Irene Moore School of Nursing
University of California, Davis
2570 48th St
Sacramento, CA, 95817
United States
Phone: 1 916 734 8558
Email: hmyoung@ucdavis.edu

Abstract

Background: Type 2 diabetes is a growing public health problem amenable to prevention and health promotion. As healthy behaviors have an impact on disease outcomes, approaches to support and sustain diabetes self-management are vital.

Objective: This study aimed to evaluate the effectiveness of a nurse coaching program using motivational interviewing paired with mobile health (mHealth) technology on diabetes self-efficacy and self-management for persons with type 2 diabetes.

Methods: This randomized controlled trial compared usual care with an intervention that entailed nurse health coaching and mHealth technology to track patient-generated health data and integrate these data into an electronic health record. The inclusion criteria were as follows: (1) enrolled at 1 of 3 primary care clinics, (2) aged 18 years or above, (3) living with type 2 diabetes, and (4) English-speaking. We collected outcome measures at baseline, 3 months, and 9 months. The primary outcome was diabetes self-efficacy; secondary outcomes were depressive symptoms, perceived stress, physical functioning, and emotional distress and anxiety. Linear regression mixed modeling estimated the population trends and individual differences in change.

Results: We enrolled 319 participants; 287 participants completed the study (155 control and 132 intervention). The participants in the intervention group had significant improvements in diabetes self-efficacy (Diabetes Empowerment Scale, 0.34; 95% CI –0.15,0.53; $P<.01$) and a decrease in depressive symptoms compared with usual care at 3 months (Patient Health Questionnaire-9; 0.89; 95% CI 0.01-1.77; $P=.05$), with no differences in the other outcomes. The differences in self-efficacy and depression scores between the 2 arms at 9 months were not sustained. The participants in the intervention group demonstrated a significant increase in physical activity (from 23,770 steps per week to 39,167 steps per week at 3 months and 32,601 per week at 9 months).

Conclusions: We demonstrated the short-term effectiveness of this intervention; however, by 9 months, although physical activity remained above the baseline, the improvements in self-efficacy were not sustained. Further research should evaluate the minimum dose of coaching required to continue progress after active intervention and the potential of technology to provide effective ongoing automated reinforcement for behavior change.

Trial Registration: ClinicalTrials.gov NCT02672176; https://clinicaltrials.gov/ct2/show/NCT02672176

(JMIR Mhealth Uhealth 2020;8(3):e16665) doi:10.2196/16665

KEYWORDS

mHealth; electronic health record; type 2 diabetes; motivational interviewing

Introduction

Background

Type 2 diabetes is a growing public health problem amenable to prevention and health promotion [1]. The prevalence of diabetes in the United States will increase from 9.3% in 2012 to an estimated 25% to 28% by 2050, with type 2 diabetes accounting for 90% to 95% of cases [2]. Physical inactivity, poor eating habits, obesity, and smoking are common risk factors for type 2 diabetes. The connections between health behavior and disease outcomes indicate the importance of a patient-centered, proactive, and evidence-based approach to prioritizing and enacting lifestyle choices [3].

Having a chronic condition has implications for all aspects of daily life as the individual navigates choices about nutrition, physical activity, sleep, stress management, and medication regimen. Bandura and Adams [4] established that self-efficacy, the belief in one's ability to influence events, can effect changes in behavior. For decades, researchers and clinicians have recognized the importance of perceived self-efficacy in the management of diabetes, including the ability of individuals to make healthy lifestyle decisions, adhere to medication and treatment regimens, and manage stress [5]. Encounters with health care providers are episodic and usually focus on monitoring and adjusting medical treatment. As optimal health in diabetes requires a more active approach by individuals to self-manage their condition and to engage in lifestyle behavioral changes, health care providers could contribute to better outcomes by offering personalized support.

Diabetes education programs and group classes may be effective in the short term but appear to be insufficient to sustain behavioral changes (eg, improvements in physical activity and healthy eating) and self-management skills [6,7]. Motivational interviewing (MI) and health coaching have the potential to customize strategies according to the individual's priorities and interests. MI is a counseling approach to build capacity to solve problems, improve self-efficacy, and support behavioral change in diabetes management [8-11]. Health coaching utilizes MI concepts to facilitate behavior change by encouraging individuals to establish attainable personal goals, brainstorm strategies to achieve goals, and self-monitor behaviors, all within the context of an interpersonal relationship with a coach [12,13]. The results of a systematic review on health coaching found improved physiological, behavioral, psychological, and social outcomes in people with chronic conditions [14]. Qualitative exploration of patient perspectives on unmet needs in self-management revealed gaps in the existing programs in their ability to support emotional regulation, psychological adjustment, and behavior change [15]. Our group previously demonstrated the effectiveness of MI and health coaching in sustaining diabetes self-efficacy in rural communities [16]. These interventions typically rely on self-report of lifestyle changes, such as diet or physical activity modifications, limiting precision in quantifying behavioral improvements.

The International Diabetes Federation outlines clinical guidelines for type 2 diabetes management, including educating patients and providers, setting goals for self-management of blood glucose, generating a structured profile, providing feedback to patients about their results, using these data to modify treatments, and engaging in shared decision making [17]. Despite these guidelines, a systematic review of type 2 diabetes management indicates that these principles have not been widely adopted in primary care. One of the gaps is that diabetes self-management education and support programs incorporating mobile health (mHealth) inconsistently integrate data and feedback to change treatment and support behavior change [18].

Mobile technology offers new opportunities to track health behaviors, provide reinforcement through immediate feedback about objective measures of behavior such as steps taken, and improve health outcomes in chronic diseases [19-21]. Technology-enabled diabetes self-management solutions with a feedback loop using patient-generated health data (PGHD) to tailor education and individualize feedback improved hemoglobin A_{1C} [22-25]. Wearable tracking devices and mHealth apps that capture health behaviors offer an objective view of daily activity, which was not previously available [26]. These technological advances become more salient when PGHD are part of the clinical record, incorporated into the care plan of adults with type 2 diabetes, offering access to data for clinicians and improved precision health opportunities.

Objective

This study examined the impact of a novel intervention using MI-based nurse health coaching combined with wearable activity trackers that integrate patient-generated activity data into the patient's electronic health record (EHR) to improve health among adults with type 2 diabetes. We hypothesized that individuals randomized to the intervention group would show overall improved self-efficacy compared with individuals in the usual care group.

Methods

Study Design

A detailed description of the study design was previously reported in a clinical trial protocol [27]. This was a randomized controlled trial with 2 arms: (1) usual care through primary care and (2) the Patient and Provider Engagement and Empowerment Through Technology ($P^2E^2T^2$) Program—nurse coaching paired with mobile sensor technology. The study was approved by the institutional review board of the University of California, Davis, and registered at ClinicalTrials.gov (NCT02672176).

Recruitment

We recruited participants from 2 suburban and 1 urban primary care clinic within an academic health center in Northern California. The inclusion criteria were as follows: (1) aged 18 years or above, (2) receiving care at 1 of the 3 clinics, (3) living with type 2 diabetes and having HbA_{1c} of 6.5% (48 mmol/mol) or higher, and (4) able to speak English. Participants were ineligible if they did not have access to a telephone, were not able to consent because of cognitive impairment, or were pregnant. Our power analysis determined that with 100 participants per arm, we would have 80% power to detect

differences in self-efficacy. We queried the health system EHR and diabetes registry to identify eligible participants who subsequently received mailed letters and telephone calls. Potential participants were told that the study focused on using enabling technology to support their health in diabetes. Study data were collected and managed using Research Electronic Data Capture tools hosted at the Clinical Translational Science Center at UC Davis [28,29]. REDCap (Research Electronic Data Capture) is a secure, Web-based software platform designed to support data capture for research studies, providing (1) an intuitive interface for validated data capture, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for data integration and interoperability with external sources. Following telephone consent, we used Research Electronic Data Capture to randomize participants to 1 of the 2 groups in a 1:1 ratio, stratified by clinic site, to either the usual care (control group) or the $P^2E^2T^2$ program. We used stratified block randomization to ensure a balanced number across the 2 groups within each site. The participants completed written consent during their onboarding session. Participants and research team members not involved in recruitment were blind to the randomization.

Clinical Trial of Patient and Provider Engagement and Empowerment Through Technology

Usual Care

Participants in this group received usual care through their primary care clinic. Usual care comprised standard health care visits with providers and access to classes, resources, and services (ie, diabetes management and weight loss education, electronic learning videos, and care coordination). At the orientation meeting, the study team members provided instruction on how to access these resources and services as well as how to use the health system's patient portal (MyChart).

Patient and Provider Engagement and Empowerment Through Technology Intervention Program Group

The intervention group participants received the same care through their primary care clinic and training as those receiving usual care regarding health system services and resources. In addition, the intervention included (1) nurse health coaching and (2) mHealth technology to track PGHD and integrate these data into the EHR (see Multimedia Appendix 1).

Nurse Health Coaching

The nurse health coaches for the intervention were 3 registered nurses (RNs) with experience in both health coaching and management of chronic disease. To promote fidelity to the intervention and a common approach to coaching participants, the nurses received core training in MI-based coaching and diabetes management. All the RN health coaches delivering the intervention completed the HealthSciences Institute's Registered Health Coach (RHC) and Chronic Care Professional training programs (www.healthsciences.org). A final performance evaluation using the Motivational Interviewing Treatment Integrity (MITI) 3.1.1 global scale evaluation tool confirmed health coaching competency before the receipt of the RHC certificate [30]. Nurses also completed a refresher course in diabetes management through the American Association of Diabetes Educators as well as the standard health system orientation on policies, procedures, and EHR training.

We paired each participant with a nurse health coach who delivered 6 individual sessions using a counseling style based on the concepts of MI. Sessions were structured to promote mutual goal setting, enhance self-efficacy in health behavior change, and assist individuals to derive meaning from data to reinforce choices and behaviors. Two RN researchers with nurse coaching experience in diabetes audited 8 of the 158 (5%) of the participant sessions and scored the coach using the MITI. They provided timely feedback to the coaches during weekly debriefing sessions, reviewed scores, and discussed optimization strategies by reviewing scenarios.

The participants had an in-person orientation with the nurse coach, followed by telephone sessions every 2 weeks for 3 months (6 contacts total). The initial MI session elicited motivations and set goals with tracking metrics to gauge the progress toward goals at subsequent sessions. Throughout the sessions, the coaches encouraged the participants to identify facilitators and barriers to achieving their health goals.

Mobile Health Technology and Integration of Patient-Generated Health Data Into the Electronic Health Record

We provided a wearable tracking device (initially, Basis Peak, then Garmin VivoSmart Heart Rate [HR]) to the intervention group participants. This device generated real-time information about steps taken, distance walked, active minutes, heart rate, and hours of sleep at night and synced the data to either an iPhone operating system mobile phone and/or iPod Touch. We provided the iPod Touch to participants who did not already possess this technology. We preinstalled MyFitnessPal, a mobile app, on the devices to allow participants to log and track nutritional consumption if they chose. We provided in-person or telephonic technical support to all participants—including the usual care group participants—throughout the duration of the study. We encouraged the participants to wear the activity tracker for the entire 9-month duration of the study.

PGHD were integrated into the EHR when participants performed regular synchronization of the activity tracker to their personal device. We used 2 connectors, Apple HealthKit and MyChart, to accomplish the automatic transmission of data to the EHR. We used Synopsis, a feature within the EHR, to design a single screen page of relevant PGHD along with clinically relevant data elements (ie, laboratory values and medications). In the case management module of Epic Electronic Health Record, we designed a summary documentation form for the nurse coaching sessions. We sent a final summary of goals and achievements to the primary care providers. Using these tools, the participants, providers, and nurse health coaches could view trends in activity levels, sleep, and nutritional intake on either their smart device or on a computer.

Changes After Trial Commencement

Early in the intervention period, we experienced an unexpected recall of the Basis Peak activity tracking device because of a

safety issue that required identifying and selecting a replacement device. The study team, in collaboration with the advisory boards, worked diligently and promptly to identify, test, and select a replacement (Garmin VivoSmart HR) and then distribute the new device to the participants in the intervention arm of the study, providing technical support to these participants as needed. This recall affected 79 participants; most of these participants received and were oriented to their new devices within 2 weeks of the recall.

Measures

The participants completed Web-based surveys at baseline, 3 months (coinciding with the end of the intervention or 3 months from baseline), and 9 months. The baseline survey included demographic information (age, gender, race and ethnicity, education, income, and insurance type), health information (common chronic illness and health status), and technology use and adoption information. Readiness to change was assessed with 2 items measuring intention, *I am intending to make changes in my diabetes self-care in the next 6 months* and *I am intending to make changes in my diabetes self-care in the next month*, and then categorized into 3 groups: precontemplators (do not intend to make changes), contemplators (intend to act in 6 months), and preparers (intend to make a change in the next month) [31]. Surveys at all 3 time points assessed study outcome measures.

Primary Study Outcome

Diabetes self-efficacy (Diabetes Empowerment Scale [DES]–Short Form) [32] is an 8-item Likert-scale survey instrument that measures diabetes-related psychosocial self-efficacy. The overall score is the sum of scores of the 8 questions in the survey, with higher scores indicating greater self-efficacy.

Secondary Outcomes

Depression severity (Patient Health Questionnaire-9 (PHQ-9) [33] is a 9-question validated survey that measures the presence and severity of depression. The score is a sum of all the responses and ranges from 0 to 27. A score of 10 or above suggests the presence of depression.

Other Measures

The surveys also included Patient-Reported Outcomes Measurement Information System (PROMIS) [34] measures (physical function and emotional distress and anxiety) and the Perceived Stress Scale (PSS) [35]. The PROMIS physical function 4-item instrument assesses the current physical function in the individual. The PROMIS emotional distress and anxiety 4-item instrument measures self-reported fear, anxious misery, and hyperarousal symptoms. The PSS is a 4-item instrument that measures the degree to which situations in one's life are determined as stressful. We evaluated physical activity data measured as steps in the intervention group (who had the activity tracker). We audited the use of MyChart features for all participants. For the intervention group, the nurse coaches recorded goals and perception of goal attainment on the part of both the participant and the nurse coach.

Statistical Methods

Descriptive analysis yielded means and SDs for continuous variables and frequencies for categorical variables. We examined distributions and collinearity to determine whether the data met the assumptions for planned statistical analyses. We compared the demographic and health-related characteristics among the individuals in the intervention group and the usual care group using Student *t* test, Wilcoxon signed rank test, chi-square test, and Fisher exact test, as appropriate. We calculated the change in outcomes over time as the difference between baseline and 3 months and baseline and 9 months for diabetes self-efficacy scores, depression severity (PHQ-9), stress score, and PROMIS measures. We used Student *t* test to compare the change in outcome between the usual care and intervention groups (significance level: $P \leq .05$). We conducted statistical analysis using Stata, version 15.0, statistical software (StataCorp, Texas, US).

In the primary analysis, we estimated the difference over time in the effects of the intervention versus usual care in the study participants, controlling for potentially relevant variables such as demographic characteristics, readiness to change, self-reported health, and comorbid disease. This was an intention-to-treat analysis with the assumption that any dropouts were missing at random. In our evaluation, we did not find any significant difference between the participants who dropped out and the participants who continued in the study. We adopted a mixed effects maximum likelihood model that accounts for the missing data from the participants who dropped out from the study. Finally, our sensitivity analysis found no difference with regard to the significance of our findings when we excluded these participants from the analysis. We included all participants, regardless of intervention completion, in the intention-to-treat analysis. We used multivariate regression modeling for all hypotheses testing to estimate within-group and across-group effects of the intervention on the outcomes (significance level: $P \leq .05$). The mixed effects models evaluated the impact of the intervention over time on the primary outcome, diabetes self-efficacy. We included a binary indicator for intervention group assignment and a group-by-time interaction term in the models to compare improvement over time between the intervention group and usual care group. We evaluated model fit using deviance tests for nested models, and the Akaike information criterion and the Bayesian information criterion for non-nested models. We assessed the estimates for the fixed effects using a predetermined significance level ($P < .05$) on 2-sided tests and 95% CIs. We used the same approach for analyzing the primary and secondary outcomes, analyzing the effect of the intervention at baseline, 3 months, and 9 months.

Results

Overview

Multimedia Appendix 2 provides a Consolidated Standards of Reporting Trials (CONSORT) flow diagram. The diabetes registry query identified 2242 potential participants. Of these, 1938 were eligible for phone recruitment. We obtained verbal consent from 392 participants by phone before randomization. A total of 319 participants attended the orientation session,

completed the written consent, and were included in the analysis. Furthermore, 32 out of 319 (10.0%) participants, 6 out of 161 (3.7%) participants from the usual care group and 26 out of 158 (16.5%) participants from the intervention group, either dropped out or were lost to follow-up over the course of the study. Of the 287 participants who completed the 9-month follow-up surveys, 155 were in the usual care group and 132 were in the intervention group. The recruitment commenced in February 2016 and the study was completed by December 2017.

Sample Characteristics

There were no significant differences in the demographics between the usual care and intervention groups (Table 1), with an almost equal distribution in gender and a mean age of 59.1 (SD 11.4) years. Most participants identified themselves as white, followed by African American, Asian, other, or more than 1 race. Furthermore, 42 out of 277 (15.2%) participants identified themselves as Hispanic and Latino. The majority of participants had attained at least some college education or completed college degrees.

Most participants were managing multiple chronic illnesses. The participants tracked metrics related to their health at varying rates as follows: blood glucose (127/319, 39.8%), laboratory results (125/319, 39.2%), physical activity (82/319, 25.7%), nutrition (65/319, 20.4%), and sleep (54/319, 16.9%). A large percentage of participants, 256 out of 319 (80.3%), had no prior experience using mobile apps or sensors. The usual care and intervention groups were similar with regard to their baseline health characteristics and experience with technology.

Table 1. Characteristics of participants in the study.

Characteristics	Total (N=319)	Control (n=161)	Intervention (n=158)	*P* value
Gender, n (%) (N=313)				.97
Female	148 (47.3)	75 (47.2)	73 (47.4)	
Male	165 (52.7)	84 (52.8)	81 (52.6)	
Age (years), mean (SD)	59.07 (11.4)	59.18 (11.5)	58.96 (11.3)	.87
Education, n (%) (N=315)				.89
High school or less	36 (11.4)	20 (12.6)	16 (10.3)	
Some college	109 (34.6)	56 (35.2)	53 (34.0)	
Associate's degree	40 (12.7)	19 (12.0)	21 (13.5)	
Bachelor's degree	63 (20.0)	29 (18.2)	34 (21.8)	
Graduate and professional degree	67 (21.3)	35 (22.0)	32 (20.5)	
Annual income, n (%) (N=276)				.06
<US $25,000	44 (15.9)	29 (21.3)	15 (10.7)	
US $25,000-US $50,000	66 (23.9)	29 (21.3)	37 (26.4)	
US $50,001-US $75,000	56 (20.3)	29 (21.3)	27 (19.3)	
US $75,001-US $100,000	47 (17.0)	17 (12.5)	30 (21.4)	
>US $100,000	63 (22.8)	32 (23.5)	31 (22.1)	
Race, n (%) (N=311)				.79
Caucasian	196 (63.0)	100 (62.9)	96 (63.2)	
African American	39 (12.5)	18 (11.3)	21 (13.8)	
Asian	27 (8.7)	16 (10.1)	11 (7.2)	
Other	30 (9.7)	14 (8.8)	16 (10.5)	
More than 1 race	19 (6.1)	11 (6.9)	8 (5.3)	
Ethnicity, n (%) (N=277)				.28
Hispanic or Latino	42 (15.2)	18 (12.9)	24 (17.5)	
Not Hispanic or Latino	235 (84.8)	122 (87.1)	113 (82.5)	
Chronic comorbidities, n (%) (N=307)				.88
No other comorbidities	115 (37.5)	55 (34.2)	60 (38.0)	
1 comorbidity	95 (30.9)	47 (29.2)	48 (30.4)	
2 comorbidities	52 (16.9)	27 (16.8)	25 (15.8)	
3 or more comorbidities	45 (14.7)	25 (15.5)	20 (12.7)	
Experience with health apps or sensors, n (%) (N=317)				.74
Previously used apps or sensors	63 (19.8)	33 (20.5)	30 (19.0)	
Never used apps or sensors	256 (80.3)	128 (79.5)	128 (81.0)	
Current health tracking, n (%)				
Blood glucose	127 (39.8)	64 (39.8)	63 (39.9)	.31
Physical activity	82 (25.7)	35 (21.7)	47 (29.8)	.25
Nutrition	65 (20.4)	33 (20.5)	32 (20.3)	.46
Sleep	54 (16.9)	28 (17.4)	26 (16.5)	.42

Intervention Engagement

The most common smart goals selected by the intervention group participants were physical activity (103/147, or 70.7%) and nutrition (37 out of 147 or 25.2%), with 7 out of 147 (4.8%) participants selecting other goals such as stress reduction, alcohol cessation, and improving sleep. Across all coaching sessions, participants averaged 172 min of nurse coaching.

Study Outcomes

Multimedia Appendices 3 and 4 summarize the descriptive results for the study outcomes. At baseline, the mean diabetes self-efficacy score was 3.66 (SD 0.89) in the usual care group and 3.67 (SD 0.83) in the intervention group. This score increased in both groups at 3 months, 3.71 (SD 0.86) in the usual care group and 4.05 (SD 0.69) in the intervention group. For the depression severity measure, PHQ-9, the usual care group experienced slightly greater depressive symptoms over time, whereas the intervention group's PHQ-9 score decreased at 3 months. There were no changes in the perceived stress or the PROMIS measures. For the intervention group at baseline, the average number of steps per week was 23,770 (SD 18,470), which increased at 3 months to 39,167 (SD 22,513) and declined at 9 months to 32,601 (SD 19,851). Furthermore, 82 of the 132 (62.1%) participants who completed the 9-month survey continued to use the device until the end of the study.

Multimedia Appendices 5 and 6 show the comparison of changes in outcomes (difference in differences) between the participants in the intervention group and the participants in the usual care group. The analysis of outcome measures at baseline and 3 months demonstrated a significant improvement in diabetes self-efficacy (DES), 0.34 (95% CI –0.15,0.53), and depressive symptoms (PHQ-9), 0.89 (95% CI 0.01-1.77), and a trend toward decreased perceived stress (0.59, 95% CI 0.03-1.16) in the intervention group compared with the usual care group. There were no significant differences in emotional distress, anxiety, or physical functioning (PROMIS) at different time intervals. There were no significant differences in the outcome measures at baseline and 9 months (end of study) between participants in the intervention and usual care groups.

Finally, our mixed effect regression models evaluated the effect of the intervention on the primary outcome, diabetes empowerment (DES), over time (3 months and 9 months). We found significant improvement in the DES scores at 3 months (0.50, 95% CI 0.07-1.1; $P<.05$) in the intervention group, but this was not sustained at 9 months after adjusting for readiness to change, self-reported health, gender, education, race, and comorbid disease in the final model. We did not find significant changes in the PHQ-9 and PSS scores after adjusting for other factors in the regression models.

Discussion

Principal Findings

This study built upon the recognition that chronic disease management is fundamentally a partnership between health care providers and individuals, requiring goal setting, bilateral communication, and motivation. We sought to change the conversation through mHealth technology and nurse coaches' support that could standardize goal setting and generate relevant patient-generated data for discussion and action. We engaged the stakeholders representing persons with diabetes, clinicians, and health information technology experts to design the intervention and the integration of the technology and coaching into primary care.

This study demonstrated the short-term effectiveness of an innovative diabetes intervention using nurse health coaching and mHealth technology on diabetes self-efficacy and increased physical activity, supporting our hypothesis. However, by 6 months post intervention, although physical activity remained above baseline, the differences in self-efficacy were not sustained.

The change in diabetes self-efficacy score represented a meaningful improvement in the confidence for engaging in self-management behavior in 3 to 4 out of 8 potential areas. The average PHQ-9 score was below the cutoff indicating substantial depressive symptoms, and the score decreased over the course of the intervention. Improved self-efficacy and management of depression may be the most important drivers for positive health behavior change. This finding also suggests a potential for PGHD derived from wearable sensors to play a role in improving self-efficacy. By using tracking devices, the participants became aware of their physical activity and behavior patterns and could visualize and measure their accomplishments, increasing motivation to continue changes to attain their goals. The nurse coaches explaining the data from sensors and supporting the participants as personal difficulties arose were critical in the positive outcomes as the coaching interactions also created accountability, focus, and awareness of how behavior impacted the participants' health.

The decrease in depression severity scores among the intervention group participants further endorse the importance of personalized support, connection to health care resources to cope with and treat depression, and continuous communication provided during the coaching sessions. Significant differences in these areas did not persist at 9 months. As lifestyle changes take time and require reinforcement, it is possible that a longer intervention period with continuing support (ie, health coaching) could be needed for sustained behavior changes leading to better diabetes management.

Participants in the intervention demonstrated a significant increase in physical activity, as measured by increased steps per week, from 23,700 to 39,167 at 3 months and down to 32,601 at 9 months. Using average stride length, this is an increase from about 11.9 miles to 19.6 miles walked per week. At 9 months, participants were still walking 16 miles. This accomplishment is consistent with the Centers for Disease Control and Prevention's (CDC's) recommended guidelines for type 2 diabetes management to get at least 150 min per week of moderate-intensity physical activity [36] and indicates an improvement in lifestyle choices. According to the CDC, incremental and sustained improvements in physical activity over time will contribute to better management of type 2 diabetes.

The lack of long-term sustainability of this intervention echoes other studies [37-39]. In a hospital-based telephone coaching

intervention for patients with type 2 diabetes, Varney et al [37] reported that the intervention was only effective during the study period but not sustainable past the completion of the study. Mohr et al [39] and Yardley et al [38] discussed the importance of human support in motivation, effective engagement, and adherence of participants to electronic health interventions.

The high retention rate (89.9% completed the 9-month study), coupled with the qualitative feedback from participants, suggests that the intervention was acceptable and useful despite the fact that over 80% had no previous experience with technology. Both groups engaged with the technology; the usual care group demonstrated interest and use of online diabetes resources, and the intervention group used their wearable activity trackers and online resources. Although the digital divide exists, this study demonstrates that mHealth solutions are viable for novices to technology and individuals with low socioeconomic status. Fu et al [40] identified specific elements that could enhance the utilization of diabetes apps, including assisting participants to recognize patterns, customizing targets, reviewing data, and planning lifestyle adjustments.

Strengths and Limitations

The passive tracking devices provided data for discussion with the nurse coaches who assisted in making sense of the patterns, generating insights into health habits that can affect outcomes that matter to the patient. These insights form the basis for long-term behavior change necessary for optimal health in diabetes. In a parallel analysis of the qualitative data from this study, participants described new insights about living with type 2 diabetes and defined success in their own terms as changes in awareness, mindset, engagement with health resources, and self-perceptions of emotional or physical health [41]. Participants in our trial substantiated the importance of context, meaning, and health care partnerships in engaging and sustaining engagement with the use of mHealth technology [19].

This study also demonstrated the feasibility of integrating such PGHD into the EHR for visualization by the patient and the health care team. Although the use of mHealth to improve diabetes self-management is well illustrated in the research literature [42-44], the data gathered in those studies were only used for research, without integration into the patient's EHR for optimal utilization by clinicians managing the disease. The EHR typically stores data generated by providers and the health system, with no contribution of data either generated by the patient or deemed important by patients in creating a complete picture of their health. To our knowledge, this is the first study reporting successful comprehensive integration of PGHD not only into the EHR at an academic medical center but also its meaningful integration into the provider's workflow.

This study had several limitations. First, the sample might have been biased toward those ready to change and those ready to use technology, limiting generalizability. Second, this study has limited generalizability to other settings because of required investments in technology, technology training, and support throughout the intervention. However, we estimated the total cost of the intervention, including staff time and technology, to be less than US $500 per participant. The total cost is a small investment relative to the costs of an office or emergency room visit. Finally, this intervention requires commitment to a model of care to include nurse coaches in a team-based approach.

Although this study demonstrated the short-term effects on self-efficacy, it did not demonstrate sustainability of the effects. Since this study commenced, 2 major trends have accelerated—the use of mHealth technology and the movement toward value-based payment for chronic disease management. These facilitating forces will likely result in approaches akin to those examined in this study. This study answered several logistical questions for scalability and generalizability, demonstrated the feasibility of integrating PGHD into the clinical record, and engaged and supported patients with little previous exposure to technology. As mobile technology is gaining adoption across all age demographics [45], scalability becomes more reasonable with more tech-savvy patients who already own mobile phones. EHR vendors are now working on HIPPA-compliant platforms that incorporate PGHD. With this forward movement in technology, the intervention becomes even more feasible.

Although traditional fee-for-service reimbursement models would not cover such interventions, managed care plans and value-based purchasing will move toward reimbursing interventions that yield quality outcomes, further enhancing the potential for sustaining interventions of this kind. Finally, as PGHD are not highly sensitive, this intervention does not have to be clinic-based. It has the potential for delivery in a variety of settings such as fitness centers, workplaces, and community centers, increasing the potential to scale.

Conclusions

Coupling patient data with other indicators of health provides a more complete picture that could be helpful in a variety of chronic conditions, such as congestive heart failure and chronic obstructive pulmonary disease, which require a partnership between clinicians and patients for effective management. This study combined nurse coaching with mHealth and compared this intervention with usual care; future studies could include more study arms that break out the components to allow greater comparison. Future research could examine this approach in additional conditions and other strategies, such as automated feedback in the form of SMS messaging and online peer support that could enhance the effectiveness of the intervention. Longer study periods with intermittent coach contact could potentially demonstrate how to sustain the effects of the intervention over time. Finally, studies of implementation and dissemination across a variety of settings would improve the translation of research such as this into practice.

Acknowledgments

The authors would like to thank the $P^2E^2T^2$ program staff Rupinder Colby, Sarina Fazio, Daicy Luo, Michael Dang, Nazifa Hamdard, and Sarah Haynes; nurse health coaches Jennifer Edwards, Bridget Levich, Katherine Greene, and Sarina Fazio; IT team members Michael Minear, Kent Anderson, Scott MacDonald, Ryan Peck, and Kristen Augtagne; primary care, health management, and education consultants Thomas Balsbaugh, Victor Baquero, Bridget Levich, Linda Blake, Glee Van Loon, and Deborah Greenwood; and the patient advisory board members Eric Bowser, Diane Goodman, Margaret Hitchcock, Maria Ibarra, Michael Lawson, Joseph McCarthy, and Tarunesh Singh for their contributions to the project.

This study was funded by the Patient-Centered Outcomes Research Institute: IHS-1310-07894. This study was conceived and designed independently of the funders, who did not have any role in the data collection, analysis, or writing of manuscript based on the data.

Authors' Contributions

All authors contributed to the design of the study, interpretation of the results, and discussion of the results. HY wrote the manuscript, with major portions contributed by SM, MD, and YF. MD conducted the statistical analysis. All authors reviewed and edited the manuscript. HY is the guarantor for this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Intervention figure.
[DOCX File , 36 KB - mhealth_v8i3e16665_app1.docx]

Multimedia Appendix 2
CONSORT flow diagram.
[DOCX File , 197 KB - mhealth_v8i3e16665_app2.docx]

Multimedia Appendix 3
Outcome measures at baseline, 3 months, and 9 months for control group table.
[DOCX File , 13 KB - mhealth_v8i3e16665_app3.docx]

Multimedia Appendix 4
Outcome measures at baseline, 3 months, and 9 months for intervention group table.
[DOCX File , 13 KB - mhealth_v8i3e16665_app4.docx]

Multimedia Appendix 5
Change in outcomes comparing baseline and 3 months (Difference in Difference).
[DOCX File , 13 KB - mhealth_v8i3e16665_app5.docx]

Multimedia Appendix 6
Change in outcomes comparing baseline and 9 months (Difference in Difference) table.
[DOCX File , 13 KB - mhealth_v8i3e16665_app6.docx]

Multimedia Appendix 7
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 189 KB - mhealth_v8i3e16665_app7.pdf]

References

1. Centers for Disease Control and Prevention. Atlanta; 2019. Type 2 Diabetes URL: https://www.cdc.gov/diabetes/basics/type2.html [accessed 2019-08-31]
2. Narayan KV, Boyle JP, Geiss LS, Saaddine JB, Thompson TJ. Impact of recent increase in incidence on future diabetes burden: US, 2005-2050. Diabetes Care 2006 Aug 25;29(9):2114-2116. [doi: 10.2337/dc06-1136]
3. American Geriatrics Society Expert Panel on the Care of Older Adults with Multimorbidity. Patient-centered care for older adults with multiple chronic conditions: a stepwise approach from the American Geriatrics Society: American Geriatrics

Society Expert Panel on the Care of Older Adults with Multimorbidity. J Am Geriatr Soc 2012 Oct;60(10):1957-1968 [FREE Full text] [doi: 10.1111/j.1532-5415.2012.04187.x] [Medline: 22994844]
4. Bandura A, Adams NE. Analysis of self-efficacy theory of behavioral change. Cogn Ther Res 1977 Dec;1(4):287-310. [doi: 10.1007/BF01663995]
5. Anderson RM, Funnell MM, Fitzgerald JT, Merrero DG. The diabetes empowerment scale: a measure of psychosocial self-efficacy. Diabetes Care 2000;23(6):739-743. [doi: 10.2337/diacare.23.6.739] [Medline: 10840988]
6. Beck J, Greenwood DA, Blanton L, Bollinger ST, Butcher MK, Condon JE, 2017 Standards Revision Task Force. 2017 National standards for diabetes self-management education and support. Diabetes Educ 2018 Feb;44(1):35-50. [doi: 10.1177/0145721718754797] [Medline: 29346744]
7. Powers MA, Bardsley J, Cypress M, Duker P, Funnell MM, Fischl AH, et al. Diabetes self-management education and support in type 2 Diabetes. Diabetes Educ 2017 Dec;43(1):40-53. [doi: 10.1177/0145721716689694] [Medline: 28118121]
8. Soderlund PD. Effectiveness of motivational interviewing for improving physical activity self-management for adults with type 2 diabetes: A review. Chronic Illn 2018 Mar;14(1):54-68. [doi: 10.1177/1742395317699449] [Medline: 29226694]
9. Lundahl B, Moleni T, Burke BL, Butters R, Tollefson D, Butler C, et al. Motivational interviewing in medical care settings: a systematic review and meta-analysis of randomized controlled trials. Patient Educ Couns 2013 Dec;93(2):157-168. [doi: 10.1016/j.pec.2013.07.012] [Medline: 24001658]
10. West DS, DiLillo V, Bursac Z, Gore SA, Greene PG. Motivational interviewing improves weight loss in women with type 2 diabetes. Diabetes Care 2007 May;30(5):1081-1087. [doi: 10.2337/dc06-1966] [Medline: 17337504]
11. Chen SM, Creedy D, Lin H, Wollin J. Effects of motivational interviewing intervention on self-management, psychological and glycemic outcomes in type 2 diabetes: a randomized controlled trial. Int J Nurs Stud 2012 Jul;49(6):637-644. [doi: 10.1016/j.ijnurstu.2011.11.011] [Medline: 22209215]
12. Wolever RQ, Simmons LA, Sforzo GA, Dill D, Kaye M, Bechard EM, et al. A systematic review of the literature on health and wellness coaching: defining a key behavioral intervention in healthcare. Glob Adv Health Med 2013 Jul;2(4):38-57 [FREE Full text] [doi: 10.7453/gahmj.2013.042] [Medline: 24416684]
13. Wolever RQ, Dreusicke M, Fikkan J, Hawkins TV, Yeung S, Wakefield J, et al. Integrative health coaching for patients with type 2 diabetes: a randomized clinical trial. Diabetes Educ 2010 Jun;36(4):629-639. [doi: 10.1177/0145721710371523] [Medline: 20534872]
14. Kivelä K, Elo S, Kyngäs H, Kääriäinen M. The effects of health coaching on adult patients with chronic diseases: a systematic review. Patient Educ Couns 2014 Dec;97(2):147-157. [doi: 10.1016/j.pec.2014.07.026] [Medline: 25127667]
15. Pal K, Dack C, Ross J, Michie S, May C, Stevenson F, et al. Health interventions for adults with type 2 diabetes: qualitative study of patient perspectives on diabetes self-management education and support. J Med Internet Res 2018 Jan 29;20(2):e40 [FREE Full text] [doi: 10.2196/jmir.8439] [Medline: 29463488]
16. Young H, Miyamoto S, Ward D, Dharmar M, Tang-Feldman Y, Berglund L. Sustained effects of a nurse coaching intervention via telehealth to improve health behavior change in diabetes. Telemed J E Health 2014 Oct;20(9):828-834 [FREE Full text] [doi: 10.1089/tmj.2013.0326] [Medline: 25061688]
17. International Diabetes Federation. Brussels: International Diabetes Federation; 2017. Managing Type 2 Diabetes in Primary Care URL: https://www.idf.org/our-activities/care-prevention/type-2-diabetes.html [accessed 2018-06-15]
18. Greenwood DA, Young HM, Quinn CC. Telehealth remote monitoring systematic review: structured self-monitoring of blood glucose and impact on A1C. J Diabetes Sci Technol 2014;8(2):378-389 [FREE Full text] [doi: 10.1177/1932296813519311] [Medline: 24876591]
19. Miyamoto SW, Henderson S, Young HM, Pande A, Han JJ. Tracking health data is not enough: a qualitative exploration of the role of healthcare partnerships and mHealth technology to promote physical activity and to sustain behavior change. JMIR Mhealth Uhealth 2016 Jan 20;4(1):e5 [FREE Full text] [doi: 10.2196/mhealth.4814] [Medline: 26792225]
20. Fatehi F, Gray LC, Russell AW. Mobile health (mHealth) for diabetes care: opportunities and challenges. Diabetes Technol Ther 2017 Dec;19(1):1-3. [doi: 10.1089/dia.2016.0430] [Medline: 28099051]
21. Berman MA, Guthrie NL, Edwards KL, Appelbaum KJ, Njike VY, Eisenberg DM, et al. Change in glycemic control with use of a digital therapeutic in adults with type 2 diabetes: cohort study. JMIR Diabetes 2018 Mar 14;3(1):e4 [FREE Full text] [doi: 10.2196/diabetes.9591] [Medline: 30291074]
22. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-management education and support. J Diabetes Sci Technol 2017 May 31;11(5):1015-1027. [doi: 10.1177/1932296817713506] [Medline: 28560898]
23. Bonoto BC, de Araújo VE, Godói IP, de Lemos LL, Godman B, Bennie M, et al. Efficacy of mobile apps to support the care of patients with diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2017 Mar 01;5(3):e4 [FREE Full text] [doi: 10.2196/mhealth.6309] [Medline: 28249834]
24. Pludwinski S, Ahmad F, Wayne N, Ritvo P. Participant experiences in a smartphone-based health coaching intervention for type 2 diabetes: a qualitative inquiry. J Telemed Telecare 2016 May;22(3):172-178. [doi: 10.1177/1357633X15595178] [Medline: 26199275]

25. Nelson LA, Coston TD, Cherrington AL, Osborn CY. Patterns of user engagement with mobile- and web-delivered self-care interventions for adults with T2DM: a review of the literature. Curr Diab Rep 2016 Jul;16(7):66 [FREE Full text] [doi: 10.1007/s11892-016-0755-1] [Medline: 27255269]
26. Gay V, Leijdekkers P. Bringing health and fitness data together for connected health care: mobile apps as enablers of interoperability. J Med Internet Res 2015;17(11):e260 [FREE Full text] [doi: 10.2196/jmir.5094] [Medline: 26581920]
27. Miyamoto S, Dharmar M, Fazio S, Tang-Feldman Y, Young HM. mHealth technology and nurse health coaching to improve health in diabetes: protocol for a randomized controlled trial. JMIR Res Protoc 2018 Feb 15;7(2):e45 [FREE Full text] [doi: 10.2196/resprot.9168] [Medline: 29449200]
28. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
29. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
30. Moyers TB, Martin T, Manuel JK, Miller WR, Ernst D. Center on Alcoholism, Substance Abuse and Addictions (CASAA). Albuquerque: Center on Alcoholism, Substance Abuse and Addictions (CASAA); 2010. Revised Global Scales: Motivational Interviewing Treatment Integrity 3.1.1 (MITI 3.1.1) URL: https://casaa.unm.edu/download/miti3_1.pdf [accessed 2016-01-08]
31. Vallis M, Ruggiero L, Greene G, Jones H, Zinman B, Rossi S, et al. Stages of change for healthy eating in diabetes: relation to demographic, eating-related, health care utilization, and psychosocial factors. Diabetes Care 2003 May;26(5):1468-1474. [doi: 10.2337/diacare.26.5.1468] [Medline: 12716806]
32. Anderson RM, Fitzgerald JT, Gruppen LD, Funnell MM, Oh MS. The Diabetes Empowerment Scale-Short Form (DES-SF). Diabetes Care 2003 May;26(5):1641-1642. [doi: 10.2337/diacare.26.5.1641-a] [Medline: 12716841]
33. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Oct;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
34. Broderick JE, DeWitt EM, Rothrock N, Crane PK, Forrest CB. Advances in patient-reported outcomes: the NIH PROMIS(®) measures. EGEMS (Wash DC) 2013;1(1):1015 [FREE Full text] [doi: 10.13063/2327-9214.1015] [Medline: 25848562]
35. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385. [doi: 10.2307/2136404]
36. Centers for Disease Control and Prevention. 2018. Living with Diabetes: Get Active URL: https://www.cdc.gov/diabetes/managing/active.html [accessed 2019-07-12]
37. Varney JE, Weiland TJ, Inder WJ, Jelinek GA. Effect of hospital-based telephone coaching on glycaemic control and adherence to management guidelines in type 2 diabetes, a randomised controlled trial. Intern Med J 2014 Oct;44(9):890-897. [doi: 10.1111/imj.12515] [Medline: 24963611]
38. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
39. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
40. Fu HN, Adam TJ, Konstan JA, Wolfson JA, Clancy TR, Wyman JF. Influence of patient characteristics and psychological needs on diabetes mobile app usability in adults with type 1 or type 2 diabetes: crossover randomized trial. JMIR Diabetes 2019 Apr 30;4(2):e11462 [FREE Full text] [doi: 10.2196/11462] [Medline: 31038468]
41. Fazio S, Edwards J, Miyamoto S, Henderson S, Dharmar M, Young HM. More than A1C: types of success among adults with type-2 diabetes participating in a technology-enabled nurse coaching intervention. Patient Educ Couns 2019 Jan;102(1):106-112. [doi: 10.1016/j.pec.2018.08.028] [Medline: 30172572]
42. Hood M, Wilson R, Corsica J, Bradley L, Chirinos D, Vivo A. What do we know about mobile applications for diabetes self-management? A review of reviews. J Behav Med 2016 Dec;39(6):981-994. [doi: 10.1007/s10865-016-9765-3] [Medline: 27412774]
43. Wu Y, Yao X, Vespasiani G, Nicolucci A, Dong Y, Kwong J, et al. Mobile app-based interventions to support diabetes self-management: a systematic review of randomized controlled trials to identify functions associated with glycemic efficacy. JMIR Mhealth Uhealth 2017 Mar 14;5(3):e35 [FREE Full text] [doi: 10.2196/mhealth.6522] [Medline: 28292740]
44. Kitsiou S, Paré G, Jaana M, Gerber B. Effectiveness of mHealth interventions for patients with diabetes: an overview of systematic reviews. PLoS One 2017;12(3):e0173160 [FREE Full text] [doi: 10.1371/journal.pone.0173160] [Medline: 28249025]
45. Pew Research Center. Washington DC: Pew Research Center; 2019 Jun 12. Mobile Fact Sheet 2018 URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2019-07-14]

Abbreviations

CDC: Centers for Disease Control and Prevention
DES: Diabetes Empowerment Scale
EHR: electronic health record
mHealth: mobile health
MI: motivational interviewing
MITI: Motivational Interviewing Treatment Integrity
$P^2E^2T^2$: Patient and Provider Engagement and Empowerment Through Technology
PGHD: patient-generated health data
PHQ-9: Patient Health Questionnaire-9
PROMIS: Patient-Reported Outcomes Measurement Information System
PSS: Perceived Stress Scale
RHC: Registered Health Coach
RN: registered nurse

Edited by G Eysenbach; submitted 12.10.19; peer-reviewed by L Schwab-Reese, A Bashir, T Scott Duncan; comments to author 01.12.19; revised version received 11.12.19; accepted 15.12.19; published 02.03.20.

Please cite as:
Young HM, Miyamoto S, Dharmar M, Tang-Feldman Y
Nurse Coaching and Mobile Health Compared With Usual Care to Improve Diabetes Self-Efficacy for Persons With Type 2 Diabetes: Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(3):e16665
URL: https://mhealth.jmir.org/2020/3/e16665
doi:10.2196/16665
PMID:32130184

Original Paper

Effectiveness of Lilly Connected Care Program (LCCP) App-Based Diabetes Education for Patients With Type 2 Diabetes Treated With Insulin: Retrospective Real-World Study

Yiyu Zhang[1,2,3], MMed; Chaoyuan Liu[4], MMed; Shuoming Luo[1,2,3], MD; Jin Huang[1], PhD, RN; Xia Li[1,2,3*], MD; Zhiguang Zhou[1,2,3*], MD

[1]Department of Metabolism and Endocrinology, The Second Xiangya Hospital, Central South University, Changsha, China

[2]Key Laboratory of Diabetes Immunology, Ministry of Education, Changsha, China

[3]National Clinical Research Center for Metabolic Diseases, Changsha, China

[4]Department of Oncology, The Second Xiangya Hospital, Central South University, Changsha, China

[*]these authors contributed equally

Corresponding Author:
Zhiguang Zhou, MD
Department of Metabolism and Endocrinology
The Second Xiangya Hospital
Central South University
No 139, Renmin Road
Changsha,
China
Phone: 86 073185292154
Fax: 86 073185367220
Email: zhouzhiguang@csu.edu.cn

Abstract

Background: Diabetes poses heavy economic and social burdens worldwide. Mobile apps show great potential for diabetes self-management education. However, there is limited evidence for the effectiveness of providing general diabetes education through mobile apps.

Objective: The aim of this study was to clarify the effectiveness of Lilly Connected Care Program (LCCP) app-based diabetes education for glycemic control.

Methods: This retrospective cohort study included patients with diabetes recruited to the LCCP platform from September 1, 2018, to May 31, 2019. Each patient was followed for 12 weeks. According to the number of diabetes education courses they had completed, the patients were divided into the following three groups: group A (0-4 courses), group B (5-29 courses), and group C (≥30 courses). The main outcomes were the change in blood glucose at the 12th week compared with baseline and the differences in blood glucose at the 12th week among the three groups. The associations of the number of diabetes education courses completed with the average blood glucose and frequency of self-monitoring of blood glucose (SMBG) at the 12th week were assessed by multivariate linear regression analyses controlling for other confounding covariates. Univariate and multivariate linear regression analyses were used to assess factors influencing patients' engagement in the diabetes education courses.

Results: A total of 5011 participants were enrolled. Their mean fasting blood glucose (FBG) and postprandial blood glucose (PBG) were significantly lower at the 12th week than at baseline (FBG, 7.46 [standard deviation (SD) 1.95] vs 7.79 [SD 2.18] mmol/L, P<.001; PBG, 8.94 [SD 2.74] vs 9.53 [SD 2.81] mmol/L, P<.001). The groups that completed more diabetes education courses had lower FBG (group B, β=−0.14, 95% CI −0.26 to −0.03; group C, β=−0.29, 95% CI −0.41 to −0.16; P for trend <.001) and PBG (group B, β=−0.29, 95% CI −0.46 to −0.11; group C, β=−0.47, 95% CI −0.66 to −0.28; P for trend <.001) and a higher frequency of SMBG at the 12th week (group B, β=1.17, 95% CI 0.81-1.53; group C, β=4.21, 95% CI 3.81-4.62; P for trend <.001) when compared with the findings in group A. Age and education were related to patients' engagement in the diabetes education courses. Middle-aged patients (35-59 years old) and elderly patients (≥60 years old) completed more diabetes education courses (middle-aged group, β=2.22, P=.01; elderly group, β=2.42, P=.02) than young patients (18-34 years old).

Conclusions: LCCP app-based diabetes education is effective for glycemic control and SMBG behavior improvement in patients with type 2 diabetes receiving insulin therapy. Young patients' engagement in the education courses was relatively low. We need to conduct in-depth interviews with users to further improve the curriculum.

(JMIR Mhealth Uhealth 2020;8(3):e17455) doi:10.2196/17455

KEYWORDS

diabetes mellitus; mobile app; diabetes self-management education; glycemic control

Introduction

Background

The prevalence of diabetes has been increasing worldwide [1,2]. It was estimated that there were 450 million adults with diabetes in 2017, and this figure is likely to increase to 690 million by 2045 [3]. According to a national survey in 2010, the prevalence of adult diabetes in China was 11.6%, representing more than 100 million adult patients with diabetes in China. Only 39.7% of those treated had ideal glycemic control [4]. Poor glycemic control can cause various complications, such as blindness, renal failure, and myocardial infarction [5]. In 2017, approximately 5 million adults died of diabetes [3]. In 2015, the global cost for diabetes was estimated to be approximately US $1.31 trillion [6].

Diabetes self-management and education are the cornerstones for diabetes management [7]. Studies have revealed that diabetes self-management education can help patients improve their glycemic control and self-management ability [8-10]. Diabetes guidelines also emphasize the importance of diabetes education [11,12]. In China, diabetes education traditionally takes the forms of group education classes in hospitals and individualized education in outpatient clinics. However, some patients do not have information on such classes, and it is inconvenient for them to take those classes because of distance and time constraints [13]. In addition, because medical resources in China are imbalanced, qualified diabetes educators are in short supply in rural areas. It may be difficult to provide high-quality diabetes education courses in primary hospitals [14]. Additionally, doctors from tertiary hospitals are overloaded with work and have limited time [15]. Outpatient consultations usually last only a few minutes. Patients receive little diabetes self-management knowledge in such a limited time [15]. Many patients with diabetes in China have not received any form of diabetes self-management education [16,17], and most Chinese patients with diabetes having poor glycemic control lack the ability to self-manage their diabetes [18].

Internet diabetes education is a potential way to overcome the barriers of distance, limited access, and short supply of qualified diabetes educators. Studies have also shown that internet-based diabetes education can improve glycemic control among patients with diabetes [19]. Mobile apps can receive and transmit information at any time and place. Compared with computers, mobile phones are easier to carry and easier to operate for the elderly [20]. With the popularity of smart phones, mobile apps have great potential for diabetes management. Many diabetes management apps provide general diabetes education for patients [21,22], and qualitative research has shown that for patients, gaining diabetes knowledge through mobile apps is more acceptable than receiving diabetes education classes in hospitals [23]. Studies have also suggested that diabetes management apps are beneficial for glycemic control and the self-management ability of patients with diabetes [24,25]. However, most studies have examined a mixture of remote monitoring, diabetes education, feedback from health care professionals, automatic feedback according to artificial intelligence, etc [21,26]. It is difficult to determine the functions of each feature [25,27,28]. Although feedback from health care professionals is important for the effectiveness of diabetes management apps [24,29,30], health care professionals are short on resources, and it is difficult to maintain their enthusiasm for using apps to manage patients, which are costly, without subsidies.

General diabetes education provided on mobile apps to patients with diabetes can make up for the shortage of diabetes educators to a certain degree. However, there is limited evidence that general diabetes education through mobile apps is effective for glycemic control among patients with diabetes. Studies have shown that the usage of diabetes apps varies across patients with different ages, education levels, and disease durations [31-33]. However, among patients who have already used apps to increase their diabetes knowledge, there is no relevant report on the difference in engagement across patients. Previous studies involving diabetes management apps were mostly randomized controlled trials with small samples. There is little real-world research on the effectiveness of diabetes management apps based on a large sample.

The Lilly Connected Care Program (LCCP) is a national diabetes care and support program that aims to improve diabetes management through internet technology and smart blood glucose monitoring devices with mobile communication for diabetes education and services in China. The LCCP is delivered via its official account on China's largest social app WeChat. Patients can record their blood glucose levels, view their historical blood glucose records, and engage in diabetes education courses on the LCCP platform.

Objectives

The aim of this study was to clarify the effectiveness of the LCCP app-based diabetes education program for glycemic control among patients with diabetes treated with insulin and to understand the factors associated with patients' engagement in diabetes education on the LCCP platform.

Methods

Design and Sample

This was a retrospective cohort study that included patients with diabetes recruited to the LCCP platform from September 1, 2018, to May 31, 2019. Randomly selected outpatients with diabetes receiving insulin therapy (with or without oral hypoglycemic agents) from most major cities of 31 provinces in mainland China were encouraged by their physicians to register on the LCCP platform without any financial incentives. After informed consent was obtained, patient demographic information, such as gender, age, education, type of diabetes, insulin regimen, and duration of diabetes, was collected. Each patient was followed up for 12 weeks. Eligible participants were patients with fasting blood glucose (FBG) and postprandial blood glucose (PBG) records on the platform at least once a week at week 1 and week 12. Patients with type 1 diabetes, patients aged <18 years, and patients with missing data on gender, age, education, and duration of diabetes were excluded from the study.

Intervention

There are 60 diabetes education courses on the LCCP platform. Patients can choose the education courses of interest on the LCCP platform to learn. Each course includes 1-5 sections, and each section takes 5-10 minutes to complete. The courses are presented in the form of audio and text. There is a small quiz after the completion of each section to test and consolidate patients' diabetes knowledge. Only when all the questions are answered correctly can the patient view the next section. Completing all 60 courses requires about 13 hours in total. The courses on the LCCP platform were created by experts in accordance with the standards of medical care for type 2 diabetes in China [12]. The courses cover patients' self-care behaviors according to the American Association of Diabetes Educators 7 Standard of Care [34], including eating healthy, being active, monitoring glucose, taking medications, solving problems, coping in healthy ways, and reducing risks. These courses enable patients to fully understand the necessity of glycemic control and the harm of diabetic complications. In addition to providing knowledge, self-management behavior change strategies are included (see Multimedia Appendix 1 for more details).

Outcome Measurements

Recruited patients were provided with a blood glucose monitoring kit that included a free intelligent glucometer and test strips. Self-monitoring of finger-prick capillary blood glucose was tested according to the glucose dehydrogenase method using an intelligent glucometer (Bionime Biotechnology [Ping Tan] Co, Ltd, Fuzhou City, China). The patients were taught to measure their FBG and PPG correctly to reduce subject bias. Data regarding patients' self-monitoring of blood glucose (SMBG) was automatically uploaded to the app platform through mobile 3G signals. The coefficient of variation of the measurement was below 5%, and the accuracy was in accordance with ISO 15197:2013 [35]. The baseline FBG and PBG were defined as the mean FBG and the mean PBG at the first week after recruitment. According to the number of diabetes education courses completed, the patients were divided into the following three groups: group A (0-4 courses), group B (5-29 courses), and group C (≥30 courses). Because the number of patients who did not complete any education courses was very small, we grouped the patients who completed 0 to 4 courses into one group and considered these patients as having the least engagement in the diabetes education courses. The remaining patients were divided into two groups according to whether they completed more than half of the total courses. The main outcomes were the change in the mean FBG and mean PBG at the 12th week compared with baseline and the difference in the mean FBG and mean PBG at the 12th week among the three groups. The secondary outcomes were the relationships between the number of diabetes education courses completed and the frequency of SMBG, as well as the factors associated with patients' engagement in the diabetes education courses. We defined patients' engagement in the diabetes education courses as the number of diabetes education courses that the patients completed.

Ethics

All patients provided written informed consent when recruited to the LCCP platform. This study was approved by the ethics committee of the Second Xiangya Hospital, Central South University.

Statistical Analysis

All continuous variables with a normal distribution are presented as means (standard deviations [SDs]). Variables with a nonnormal distribution are presented as medians (IQRs). Categorical variables are presented as the frequency (number of cases [n]) and percentage (%) of total study patients. A paired *t*-test was used to assess the change in blood glucose from baseline to week 12. For intragroup comparisons, an analysis of variance (ANOVA) test was used. The associations of the number of diabetes education courses completed with the average blood glucose and SMBG at the 12th week were assessed using multivariate linear regression analysis while controlling for other confounding covariates. To evaluate linear trends, we entered the median level of the diabetes education courses completed by category into the model as a continuous variable [36-38]. Univariate and multivariate linear regression analyses were used to assess the factors influencing patients' engagement in the diabetes education courses. Statistical analyses were performed using SAS 9.4 software (SAS Institute, Cary, North Carolina) via SAS Enterprise Guide version 7.1. A P value and P for trend ≤.05 were considered statistically significant.

Results

Patient Characteristics

From September 1, 2018, to May 31, 2019, a total of 5011 patients with type 2 diabetes who were older than 18 years and were receiving insulin therapy were enrolled in the study. The patient inclusion flow chart is shown in Figure 1. The sample was recruited from 31 provinces across China. Among the 5011 patients, 56.02% (2807/5011) were male, and the median age was 52.0 years (IQR 43.0-60.0 years). The median duration of

diabetes was 2.25 years (IQR 0.08-9.50 years). Patient characteristics at baseline are shown in Table 1.

Effect on Glycemic Control

The mean FBG at baseline was 7.79 (SD 2.18) mmol/L, and the mean PBG was 9.53 (SD 2.81) mmol/L. The mean FBG and PBG of the patients were significantly lower at the 12th week than at baseline (FBG, 7.46 [SD 1.95] vs 7.79 [SD 2.18] mmol/L, *P*<.001; PBG, 8.94 [SD 2.74] vs 9.53 [SD 2.81] mmol/L, *P*<.001). Among the patients, 63.10% (3162/5011) had poor baseline glycemic control (FBG ≥7 mmol/L or PBG ≥11 mmol/L). Among the patients with poor baseline blood glucose values, both FBG and PBG at the 12th week were significantly decreased from baseline (FBG, 8.02 [SD 2.04] vs 8.83 [SD 2.07] mmol/L, *P*<.001; PBG, 9.48 [SD 2.93] vs 10.52 [SD 2.93] mmol/L, *P*<.001).

Figure 1. The patient inclusion flow chart. FBG: fasting blood glucose; LCCP: Lilly Connected Care Program; PBG: postprandial blood glucose.

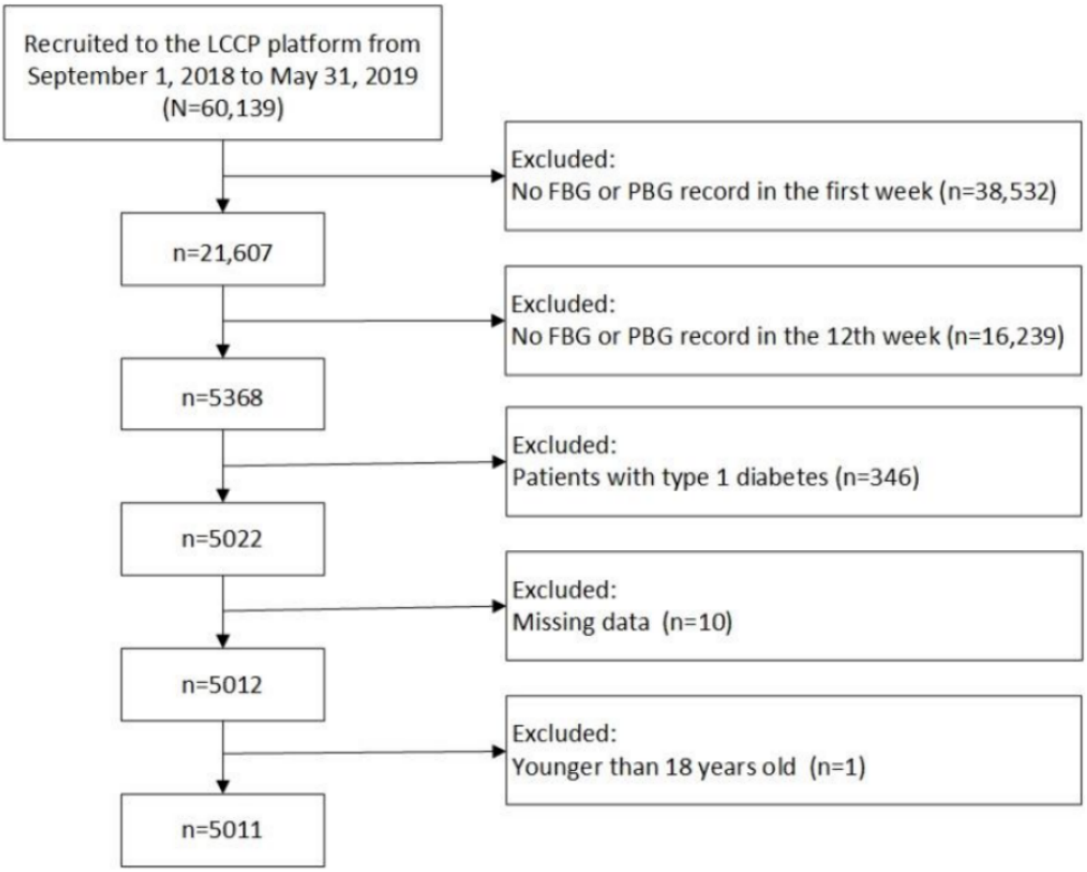


Table 1. Patient characteristics at baseline (N=5011).

Variable	Value[a]
Duration of diabetes (years)	2.25 (0.08-9.50)
Age (years)	52.0 (43.0-60.0)
Age group (years)	
18-34	480 (9.58)
35-59	3157 (63.00)
≥60	1374 (27.42)
Gender	
Male	2807 (56.02)
Female	2204 (43.98)
Education	
Junior middle school or below	1372 (27.38)
High school	1644 (32.81)
College or above	1995 (39.81)
Insulin regimen	
Premixed insulin	3909 (78.0)
Fast-acting insulin (with/without long-acting insulin)	1102 (21.99)
Region	
Northeast China	963 (19.22)
North China	972 (19.40)
East China	1395 (27.84)
South China	400 (7.98)
Central China	733 (14.63)
Northwest China	248 (4.95)
Southwest China	300 (5.99)

[a]Continuous data are presented as medians (interquartile ranges), and categorical data are presented as n (%).

According to the number of diabetes education courses completed, the patients were divided into the following three groups: group A (0-4 courses), group B (5-29 courses), and group C (≥30 courses). The mean FBG and PBG of each group decreased compared with baseline (all *P*<.001; Table 2).

Among the three groups, the mean FBG and PBG were lower in group C than in group B and group A at the 12th week (FBG, 7.28 [SD 1.87] vs 7.44 [SD 1.91] vs 7.67 [SD 2.08] mmol/L, *P*<.001; PBG, 8.70 [SD 2.54] vs 8.91 [SD 2.65] vs 9.26 [SD 3.06] mmol/L, *P*<.001). After adjusting for multiple confounding factors, including gender, age, education, duration of diabetes, baseline insulin regimen, and baseline FBG and PBG, multivariate linear regression analysis showed that the number of diabetes education courses completed was still related to both FBG and PBG (Table 3). Patients in groups B and C, who completed more diabetes education courses, had lower mean FBG (group B, β=−0.14, 95% CI −0.26 to −0.03; group C, β=−0.29, 95% CI −0.41 to −0.16; *P* for trend <.001) and PBG at the 12th week (group B, β=−0.29, 95% CI −0.46 to −0.11; group C, β=−0.47, 95% CI −0.66 to −0.28; *P* for trend <.001) compared with the findings for patients in group A (the lowest number of diabetes education courses completed).

Table 2. Comparison of mean blood glucose at week 12 and baseline.

Group	Participants, n (%)	Baseline FBG[a] (mmol/L), mean (SD)	FBG (mmol/L), mean (SD)	*P* value	Baseline PBG[b] (mmol/L), mean (SD)	PBG (mmol/L), mean (SD)	*P* value
Total	5011 (100)	7.79 (2.18)	7.46 (1.95)	<.001	9.53 (2.81)	8.94 (2.74)	<.001
Group A	1328 (26.5)	7.93 (2.26)	7.67 (2.08)	<.001	9.64 (3.08)	9.26 (3.06)	<.001
Group B	2258 (45.1)	7.76 (2.15)	7.44 (1.91)	<.001	9.56 (2.75)	8.91 (2.65)	<.001
Group C	1425 (28.4)	7.69 (2.15)	7.28 (1.87)	<.001	9.37 (2.64)	8.70 (2.54)	<.001

[a]FBG: fasting blood glucose.

[b]PBG: postprandial blood glucose.

Table 3. The relationship between mean blood glucose at the 12th week and the number of diabetes education courses completed according to linear regression analysis.

Outcome and group	Crude model			Adjusted model[a]		
	β (95% CI)	*P* value	*P* value for trend	β (95% CI)	*P* value	*P* value for trend
FBG[b] (12th week)						
Group A[c]	—[d]	—	<.001	—	—	<.001
Group B	−0.23 (−0.36 to −0.10)	<.001		−0.14 (−0.26 to −0.03)	.01	
Group C	−0.39 (−0.54 to −0.25)	<.001		−0.29 (−0.41 to −0.16)	<.001	
PBG[e] (12th week)						
Group A[c]	—	—	<.001	—	—	<.001
Group B	−0.34 (−0.53 to −0.16)	<.001		−0.29 (−0.46 to −0.11)	.001	
Group C	−0.56 (−0.76 to −0.35)	<.001		−0.47 (−0.66 to −0.28)	<.001	

[a]Adjusted for gender, age, education, duration of diabetes, baseline insulin regimen, and baseline FBG and PBG.

[b]FBG: fasting blood glucose.

[c]Reference group.

[d]Not applicable.

[e]PBG: postprandial blood glucose.

Effect on Self-Monitoring of Blood Glucose Behavior

The mean frequency of SMBG at the 12th week was 7.95 (SD 5.79) times per week. Among the three groups, the frequency of SMBG was higher in group C than in group B and group A at the 12th week (10.97 [SD 7.06] vs 7.39 [SD 4.92] vs 5.67 [SD 4.00] times per week, *P*<.001). After adjusting for multiple confounding factors, including gender, age, education, duration of diabetes, baseline FBG and PBG, and baseline SMBG frequency, multivariate linear regression analysis showed that the number of diabetes education courses that patients completed was still related to the frequency of SMBG at the 12th week (Table 4). Patients in groups B and C, who completed more diabetes education courses, had higher frequencies of SMBG at the 12th week (group B, β=1.17, 95% CI 0.81-1.53; group C, β=4.21, 95% CI 3.81-4.62; *P* for trend <.001) compared with the findings for patients in group A (the lowest number of diabetes education courses completed).

Table 4. The relationship between self-monitoring of blood glucose frequency at the 12th week and the number of diabetes education courses completed according to linear regression analysis.

Group	SMBG[a] frequency (times per week), mean (SD)	Crude model			Adjusted model[b]		
		β (95% CI)	*P* value	*P* value for trend	β (95% CI)	*P* value	*P* value for trend
Group A[c]	5.67 (4.01)	—[d]	—	<.001	—	—	<.001
Group B	7.39 (4.92)	1.72 (1.35-2.09)	<.001		1.17 (0.81-1.53)	<.001	
Group C	10.97 (7.06)	5.30 (4.9-5.71)	<.001		4.21 (3.81-4.62)	<.001	

[a]SMBG: self-monitoring of blood glucose.

[b]Adjusted for gender, age, education level, duration of diabetes, baseline SMBG frequency, fasting blood glucose, and postprandial blood glucose.

[c]Reference group.

[d]Not applicable.

Factors Influencing Engagement

The median number of diabetes education courses completed in 12 weeks was 14 (IQR 4-33). Univariate linear regression analysis showed that age (middle-aged patients, β=2.29, *P*=.01; elderly patients, β=2.72, *P*=.005), education (high school, β=1.40, *P*=.04), and baseline FBG (β=−0.27, *P*=.02) and PBG (β=−0.21, *P*=.02) were related to the number of diabetes education courses completed. Gender (*P*=.24) and duration of diabetes (*P*=.10) did not show such an association. After mutual adjustment by multivariate linear regression, the age and education level of patients were still related to the number of diabetes education courses completed (Table 5). Compared with young patients (18-34 years old), middle-aged patients (35-59 years old) and elderly patients (≥60 years old) completed more diabetes education courses (middle-aged group, β=2.22, *P*=.01; elderly group, β=2.42, *P*=.02). Compared with patients having a junior middle school education or below, the number of diabetes education courses completed was higher among patients having a high school education (β=1.46, *P*=.03) but was not significantly different among patients having a college education or above (*P*=.98).

Table 5. Factors associated with the number of diabetes education courses completed according to linear regression analysis.

Variables	Univariate model	*P* value	Multivariate model	*P* value
	β (95% CI)		β (95% CI)	
Gender				
Male[a]	—[b]	—	—	—
Female	0.60 (−0.40 to 1.61)	.24	0.61 (−0.41 to 1.62)	.24
Age (years)				
18-34[a]	—	—	—	—
35-59	2.29 (0.55 to 4.03)	.01	2.22 (0.46 to 3.98)	.01
≥60	2.72 (0.84 to 4.60)	.005	2.42 (0.43 to 4.41)	.02
Education				
Junior middle school or below[a]	—	—	—	—
High school	1.40 (0.10 to 2.70)	.04	1.46 (0.16 to 2.75)	.03
College or above	−0.26 (−1.51 to 0.98)	.68	−0.02 (−1.28 to 1.24)	.98
Duration of diabetes	0.06 (−0.01 to 0.12)	.10	0.05 (−0.02 to 0.12)	.15
Baseline FBG[c]	−0.27 (−0.50 to −0.04)	.02	−0.22 (−0.49 to 0.05)	.11
Baseline PBG[d]	−0.21 (−0.39 to −0.03)	.02	−0.18 (−0.38 to 0.03)	.10

[a]Reference group.

[b]Not applicable.

[c]FBG: fasting blood glucose.

[d]PBG: postprandial blood glucose.

Discussion

Principal Findings

Diabetes management apps are beneficial for glycemic control in patients with diabetes. However, the results of different studies showed great heterogeneity, which might be related to the different features of apps. Most studies examined a mixture of telemonitoring, education, feedback, and other functions. It was not clear which function played a role [27]. A study by Dong et al found that health education via WeChat in conjunction with conventional diabetes treatment could improve glycemic control in patients with type 2 diabetes. Because the intervention was not blinded, a certain degree of collaborative intervention might exist [39]. Our study showed that diabetes education courses on the mobile app-based LCCP platform were helpful for improving glycemic control in patients with diabetes treated with insulin. After 3 months of obtaining diabetes self-management knowledge on the LCCP platform, both the FBG and PBG of patients decreased.

The use frequency of diabetes management apps varies across different patients. Most studies did not explore the effect of app use frequency on glycemic control. A study by Agarwal et al found that a diabetes management app did not affect patients' diabetes management. A possible reason was the low use of the app among participants [40]. Vehi et al found a relevant effect of a diabetes management app on blood glucose, independent of the use frequency [41]. By recording the number of diabetes education courses that the patients completed, our study found that completion of more diabetes education courses by patients was associated with more improvement in their glycemic control. Previous studies showed that the effect of diabetes education was related to age, duration of diabetes, and education level [7,42]. After adjusting for multiple confounding factors, including gender, age, education, duration of diabetes, baseline insulin regimen, and baseline FBG and PBG, the number of diabetes education courses that the patients completed was still related to their glycemic control. A systematic review showed that diabetes self-management education can improve the glycemic control of patients, and the effect was higher for interventions that offered more than 10 contact hours [43]. This finding is consistent with our finding. A qualitative study found that existing diabetes education programs may not adequately meet all the needs of patients with type 2 diabetes [44]. Offline diabetes education classes need diabetes educator resources, and there are many barriers preventing patients from taking education classes in hospitals regularly [13,45]. No in-person diabetes education was provided in our study. Although individualized diabetes education is emphasized [46], app-based general diabetes education can save resources, is convenient for patients, and is inexpensive. It can be used as a supplement to traditional diabetes education and can mitigate the deficiencies of traditional diabetes education to some extent.

By comparing the frequency of SMBG in patients with different numbers of diabetes education courses completed, we found that the frequency of SMBG was related to the completion of diabetes education courses. After adjusting for multiple confounding factors, we found that completion of more diabetes education courses by patients was associated with a higher frequency of SMBG. Although the role of SMBG is controversial in patients with type 2 diabetes treated with oral agents, SMBG in patients with type 2 diabetes treated with insulin is very important; it is useful for adjusting insulin dosage, guiding nutrition therapy and physical activity, and preventing hypoglycemia [11,47]. The difference in the glycemic control improvement of patients with different numbers of diabetes education courses completed may partly be related to differences in their SMBG behavior. However, to be useful, SMBG information needs to be integrated with patients' clinical and self-management plans [47]. Therefore, if blood glucose monitoring data on the app platform are supplemented with feedback from health care professionals, the effect of the app might be stronger.

Previous studies showed that diabetes education can improve patients' glycemic control and self-management behaviors [7,10,48]. According to the health belief model and extended unified theory of acceptance and use of technology, patients' belief in health risk and perceived usefulness can predict the likelihood of their engagement in health behaviors [49]. The diabetes education courses on the LCCP platform emphasize the necessity of glycemic control and the harm of diabetes complications. After completing the courses, patients might realize the harm of diabetes and the usefulness of glycemic control and improve their self-care behaviors. Strategies for improving self-management behaviors are also provided in the LCCP education courses. These might be the possible mechanisms by which the educational content on the LCCP platform led to behavior changes and glycemic control improvement.

Our study not only identified the effect of diabetes education courses but also analyzed the factors influencing patients' engagement in these courses. We found that patients' age and education were related to their engagement in diabetes education courses. A previous study found that women and patients with a higher education status were more willing to participate in offline diabetes education than men and patients with a lower education status [50]. However, our study did not find gender differences. Previous studies found that health app usage was higher in younger patients than in older patients [51,52]. However, our study found that among patients who had already used the app, middle-aged and elderly patients were more involved than younger patients. The reasons for refusing to participate in diabetes education include good overall wellbeing and individuals' sense that they can manage their disease and that they have sufficient knowledge [7]. Our previous study found that effort expectancy had only a slight effect on patients' willingness to use diabetes apps [53]. With the widespread use of smart phones among the elderly, they may not experience difficulty operating such apps. Additionally, young patients may have less time to learn because they may have to work, may have other channels through which to gain diabetes knowledge, or may have a higher self-efficacy in diabetes management, which may explain why they engaged less in diabetes education courses on the app. Perceived usefulness is the most important determinant of patients' intention to use diabetes apps [54]. We should conduct in-depth interviews with

users to further identify the factors that affect patients' perceived usefulness of the education courses to further improve the curriculum.

Strengths and Limitations

Most studies on the effectiveness of diabetes apps and diabetes education programs were not blinded randomized controlled trials and had small samples. A certain degree of performance bias might be unavoidable. Our study was based on real-world data with a large sample. By comparing the differences in glycemic control between patients with different engagement levels in diabetes education courses, we found that diabetes education on the app platform was effective in diabetes management among patients with type 2 diabetes treated with insulin.

However, our study had several limitations. First, our observation period was short. The long-term effect of diabetes education provided through the app needs to be further investigated, and future studies could adopt glycosylated hemoglobin as an indicator of blood glucose levels. In addition, we did not collect other information, such as rural or urban residence, social-economic status, and chronic complications. These factors may have some impact on glycemic control and patients' participation in diabetes education courses. Finally, we found that the LCCP app-based diabetes education was associated with improvement in glycemic control, but it was unclear whether the improvement in glycemic control was due to improvements in patients' self-management behaviors. Self-management behaviors include a healthy diet, regular physical activity, SMBG, and medication adherence [34]. Our study observed improvements only in SMBG behavior; improvements in diet and exercise behaviors and drug compliance need to be further investigated.

Conclusions

LCCP app-based diabetes education is effective for glycemic control and improvement in SMBG behavior among patients with type 2 diabetes receiving insulin therapy. Young patients' engagement in these education courses is relatively low. We need to conduct in-depth interviews with users to identify the factors that affect patients' perceived usefulness of the education courses to further improve the curriculum.

Acknowledgments

The authors thank all the patients who participated in this study. The authors also thank American Journal Experts for their assistance in language editing. This work was supported by the National Key R&D Program of China (2018YFC1315603) and the Science and Technology Major Project of Hunan Province (2017SK1020).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Main contents of the Lilly Connected Care Program diabetes education courses.
[DOCX File , 13 KB - mhealth_v8i3e17455_app1.docx]

References

1. Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and Trends in Diabetes Among Adults in the United States, 1988-2012. JAMA 2015 Sep 08;314(10):1021-1029. [doi: 10.1001/jama.2015.10029] [Medline: 26348752]
2. Xu G, Liu B, Sun Y, Du Y, Snetselaar LG, Hu FB, et al. Prevalence of diagnosed type 1 and type 2 diabetes among US adults in 2016 and 2017: population based study. BMJ 2018 Sep 04;362:k1497 [FREE Full text] [doi: 10.1136/bmj.k1497] [Medline: 30181166]
3. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract 2018 Apr;138:271-281. [doi: 10.1016/j.diabres.2018.02.023] [Medline: 29496507]
4. Xu Y, Wang L, He J, Bi Y, Li M, Wang T, 2010 China Noncommunicable Disease Surveillance Group. Prevalence and control of diabetes in Chinese adults. JAMA 2013 Sep 04;310(9):948-959. [doi: 10.1001/jama.2013.168118] [Medline: 24002281]
5. Nathan DM. Long-term complications of diabetes mellitus. N Engl J Med 1993 Jun 10;328(23):1676-1685. [doi: 10.1056/NEJM199306103282306] [Medline: 8487827]
6. Bommer C, Heesemann E, Sagalova V, Manne-Goehler J, Atun R, Bärnighausen T, et al. The global economic burden of diabetes in adults aged 20-79 years: a cost-of-illness study. Lancet Diabetes Endocrinol 2017 Jun;5(6):423-430. [doi: 10.1016/S2213-8587(17)30097-9] [Medline: 28456416]
7. Świątoniowska N, Sarzyńska K, Szymańska-Chabowska A, Jankowska-Polańska B. The role of education in type 2 diabetes treatment. Diabetes Res Clin Pract 2019 May;151:237-246. [doi: 10.1016/j.diabres.2019.04.004] [Medline: 31063855]
8. Zheng F, Liu S, Liu Y, Deng L. Effects of an Outpatient Diabetes Self-Management Education on Patients with Type 2 Diabetes in China: A Randomized Controlled Trial. J Diabetes Res 2019;2019:1073131 [FREE Full text] [doi: 10.1155/2019/1073131] [Medline: 30800684]

9. Jiang X, Jiang H, Lu Y, Liu S, Wang J, Tang R, et al. The effectiveness of a self-efficacy-focused structured education programme on adults with type 2 diabetes: A multicentre randomised controlled trial. J Clin Nurs 2019 Sep;28(17-18):3299-3309. [doi: 10.1111/jocn.14908] [Medline: 31066115]
10. Hermanns N, Ehrmann D, Schall S, Maier B, Haak T, Kulzer B. The effect of an education programme (MEDIAS 2 BSC) of non-intensive insulin treatment regimens for people with Type 2 diabetes: a randomized, multi-centre trial. Diabet Med 2017 Aug;34(8):1084-1091. [doi: 10.1111/dme.13346] [Medline: 28257159]
11. American Diabetes Association. 4. Lifestyle Management: Standards of Medical Care in Diabetes-2018. Diabetes Care 2018 Jan;41(Suppl 1):S38-S50. [doi: 10.2337/dc18-S004] [Medline: 29222375]
12. Jia W, Weng J, Zhu D, Ji L, Lu J, Zhou Z, Chinese Diabetes Society. Standards of medical care for type 2 diabetes in China 2019. Diabetes Metab Res Rev 2019 Sep;35(6):e3158. [doi: 10.1002/dmrr.3158] [Medline: 30908791]
13. Mc Sharry J, Dinneen SF, Humphreys M, O'Donnell M, O'Hara MC, Smith SM, et al. Barriers and facilitators to attendance at Type 2 diabetes structured education programmes: a qualitative study of educators and attendees. Diabet Med 2019 Jan;36(1):70-79. [doi: 10.1111/dme.13805] [Medline: 30156335]
14. Chen Y, Yin Z, Xie Q. Suggestions to ameliorate the inequity in urban/rural allocation of healthcare resources in China. Int J Equity Health 2014;13:34 [FREE Full text] [doi: 10.1186/1475-9276-13-34] [Medline: 24884614]
15. Hu Y, Zhang Z. Skilled doctors in tertiary hospitals are already overworked in China. Lancet Glob Health 2015 Dec;3(12):e737 [FREE Full text] [doi: 10.1016/S2214-109X(15)00192-8] [Medline: 26566744]
16. McGuire HC, Ji L, Kissimova-Skarbek K, Whiting D, Aguirre F, Zhang P, et al. Type 1 diabetes mellitus care and education in China: The 3C study of coverage, cost, and care in Beijing and Shantou. Diabetes Res Clin Pract 2017 Jul;129:32-42. [doi: 10.1016/j.diabres.2017.02.027] [Medline: 28500868]
17. Xu H, Luo J, Wu B. Self-reported diabetes education among Chinese middle-aged and older adults with diabetes. J Glob Health 2016 Dec;6(2):020402 [FREE Full text] [doi: 10.7189/jogh.06.020402] [Medline: 27698998]
18. Zhou Y, Liao L, Sun M, He G. Self-care practices of Chinese individuals with diabetes. Exp Ther Med 2013 Apr;5(4):1137-1142 [FREE Full text] [doi: 10.3892/etm.2013.945] [Medline: 23599736]
19. Pereira K, Phillips B, Johnson C, Vorderstrasse A. Internet delivered diabetes self-management education: a review. Diabetes Technol Ther 2015 Jan;17(1):55-63. [doi: 10.1089/dia.2014.0155] [Medline: 25238257]
20. Sun C, Sun L, Xi S, Zhang H, Wang H, Feng Y, et al. Mobile Phone-Based Telemedicine Practice in Older Chinese Patients with Type 2 Diabetes Mellitus: Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Jan 04;7(1):e10664 [FREE Full text] [doi: 10.2196/10664] [Medline: 30609983]
21. Wu Y, Zhou Y, Wang X, Zhang Q, Yao X, Li X, et al. A Comparison of Functional Features in Chinese and US Mobile Apps for Diabetes Self-Management: A Systematic Search in App Stores and Content Analysis. JMIR Mhealth Uhealth 2019 Aug 28;7(8):e13971 [FREE Full text] [doi: 10.2196/13971] [Medline: 31464191]
22. Nie L, Xie B, Yang Y, Shan YM. Characteristics of Chinese m-Health Applications for Diabetes Self-Management. Telemed J E Health 2016 Jul;22(7):614-619. [doi: 10.1089/tmj.2015.0184] [Medline: 27171016]
23. Zhang Y, Li X, Luo S, Liu C, Liu F, Zhou Z. Exploration of Users' Perspectives and Needs and Design of a Type 1 Diabetes Management Mobile App: Mixed-Methods Study. JMIR Mhealth Uhealth 2018 Sep 21;6(9):e11400 [FREE Full text] [doi: 10.2196/11400] [Medline: 30249580]
24. Hou C, Carter B, Hewitt J, Francisa T, Mayor S. Do Mobile Phone Applications Improve Glycemic Control (HbA1c) in the Self-management of Diabetes? A Systematic Review, Meta-analysis, and GRADE of 14 Randomized Trials. Diabetes Care 2016 Nov;39(11):2089-2095. [doi: 10.2337/dc16-0346] [Medline: 27926892]
25. Wu X, Guo X, Zhang Z. The Efficacy of Mobile Phone Apps for Lifestyle Modification in Diabetes: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
26. Veazie S, Winchell K, Gilbert J, Paynter R, Ivlev I, Eden KB, et al. Rapid Evidence Review of Mobile Applications for Self-management of Diabetes. J Gen Intern Med 2018 Jul;33(7):1167-1176. [doi: 10.1007/s11606-018-4410-1] [Medline: 29740786]
27. Zhou W, Chen M, Yuan J, Sun Y. Welltang - A smart phone-based diabetes management application - Improves blood glucose control in Chinese people with diabetes. Diabetes Res Clin Pract 2016 Jun;116:105-110. [doi: 10.1016/j.diabres.2016.03.018] [Medline: 27321324]
28. Shan R, Sarkar S, Martin SS. Digital health technology and mobile devices for the management of diabetes mellitus: state of the art. Diabetologia 2019 Jun;62(6):877-887. [doi: 10.1007/s00125-019-4864-7] [Medline: 30963188]
29. Hou C, Xu Q, Diao S, Hewitt J, Li J, Carter B. Mobile phone applications and self-management of diabetes: A systematic review with meta-analysis, meta-regression of 21 randomized trials and GRADE. Diabetes Obes Metab 2018 Dec;20(8):2009-2013. [doi: 10.1111/dom.13307] [Medline: 29582538]
30. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A Systematic Review of Reviews Evaluating Technology-Enabled Diabetes Self-Management Education and Support. J Diabetes Sci Technol 2017 Sep;11(5):1015-1027. [doi: 10.1177/1932296817713506] [Medline: 28560898]

31. Boyle L, Grainger R, Hall RM, Krebs JD. Use of and Beliefs About Mobile Phone Apps for Diabetes Self-Management: Surveys of People in a Hospital Diabetes Clinic and Diabetes Health Professionals in New Zealand. JMIR Mhealth Uhealth 2017 Jun 30;5(6):e85 [FREE Full text] [doi: 10.2196/mhealth.7263] [Medline: 28666975]
32. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
33. Zhang Y, Li X, Luo S, Liu C, Xie Y, Guo J, et al. Use, Perspectives, and Attitudes Regarding Diabetes Management Mobile Apps Among Diabetes Patients and Diabetologists in China: National Web-Based Survey. JMIR Mhealth Uhealth 2019 Feb 08;7(2):e12658 [FREE Full text] [doi: 10.2196/12658] [Medline: 30735147]
34. AADE. AADE7 Self-Care Behaviors. Diabetes Educ 2008;34(3):445-449. [doi: 10.1177/0145721708316625] [Medline: 18535317]
35. Freckmann G, Link M, Schmid C, Pleus S, Baumstark A, Haug C. System Accuracy Evaluation of Different Blood Glucose Monitoring Systems Following ISO 15197:2013 by Using Two Different Comparison Methods. Diabetes Technol Ther 2015 Sep;17(9):635-648. [doi: 10.1089/dia.2015.0085] [Medline: 26110670]
36. Jiao J, Liu G, Shin HJ, Hu FB, Rimm EB, Rexrode KM, et al. Dietary fats and mortality among patients with type 2 diabetes: analysis in two population based cohort studies. BMJ 2019 Jul 02;366:l4009 [FREE Full text] [doi: 10.1136/bmj.l4009] [Medline: 31266749]
37. Song S, Lee JE. Dietary Patterns Related to Triglyceride and High-Density Lipoprotein Cholesterol and the Incidence of Type 2 Diabetes in Korean Men and Women. Nutrients 2018 Dec 20;11(1) [FREE Full text] [doi: 10.3390/nu11010008] [Medline: 30577478]
38. Rasmussen MG, Grøntved A, Blond K, Overvad K, Tjønneland A, Jensen MK, et al. Associations between Recreational and Commuter Cycling, Changes in Cycling, and Type 2 Diabetes Risk: A Cohort Study of Danish Men and Women. PLoS Med 2016 Jul;13(7):e1002076 [FREE Full text] [doi: 10.1371/journal.pmed.1002076] [Medline: 27403867]
39. Dong Y, Wang P, Dai Z, Liu K, Jin Y, Li A, et al. Increased self-care activities and glycemic control rate in relation to health education via Wechat among diabetes patients: A randomized clinical trial. Medicine (Baltimore) 2018 Dec;97(50):e13632 [FREE Full text] [doi: 10.1097/MD.0000000000013632] [Medline: 30558051]
40. Agarwal P, Mukerji G, Desveaux L, Ivers NM, Bhattacharyya O, Hensel JM, et al. Mobile App for Improved Self-Management of Type 2 Diabetes: Multicenter Pragmatic Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Jan 10;7(1):e10321 [FREE Full text] [doi: 10.2196/10321] [Medline: 30632972]
41. Vehi J, Regincós Isern J, Parcerisas A, Calm R, Contreras I. Impact of Use Frequency of a Mobile Diabetes Management App on Blood Glucose Control: Evaluation Study. JMIR Mhealth Uhealth 2019 Mar 07;7(3):e11933 [FREE Full text] [doi: 10.2196/11933] [Medline: 30843865]
42. Chen S, Burström B, Sparring V, Qian D, Burström K. Differential Impact of an Education-Based Intervention for Patients with Type 2 Diabetes Mellitus in Rural China. Int J Environ Res Public Health 2019 Jul 26;16(15) [FREE Full text] [doi: 10.3390/ijerph16152676] [Medline: 31357408]
43. Chrvala CA, Sherr D, Lipman RD. Diabetes self-management education for adults with type 2 diabetes mellitus: A systematic review of the effect on glycemic control. Patient Educ Couns 2016 Dec;99(6):926-943 [FREE Full text] [doi: 10.1016/j.pec.2015.11.003] [Medline: 26658704]
44. Pal K, Dack C, Ross J, Michie S, May C, Stevenson F, et al. Digital Health Interventions for Adults With Type 2 Diabetes: Qualitative Study of Patient Perspectives on Diabetes Self-Management Education and Support. J Med Internet Res 2018 Jan 29;20(2):e40 [FREE Full text] [doi: 10.2196/jmir.8439] [Medline: 29463488]
45. Preechasuk L, Sriussadaporn P, Likitmaskul S. The obstacles to diabetes self-management education and support from healthcare professionals' perspectives: a nationwide survey. Diabetes Metab Syndr Obes 2019;12:717-727 [FREE Full text] [doi: 10.2147/DMSO.S195916] [Medline: 31190931]
46. Beck J, Greenwood DA, Blanton L, Bollinger ST, Butcher MK, Condon JE, et al. 2017 National Standards for Diabetes Self-Management Education and Support. Diabetes Educ 2019 Dec;45(1):34-49. [doi: 10.1177/0145721718820941] [Medline: 30558523]
47. American Diabetes Association. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2018. Diabetes Care 2018 Jan;41(Suppl 1):S55-S64. [doi: 10.2337/dc18-S006] [Medline: 29222377]
48. Ehrmann D, Kulzer B, Schipfer M, Lippmann-Grob B, Haak T, Hermanns N. Efficacy of an Education Program for People With Diabetes and Insulin Pump Treatment (INPUT): Results From a Randomized Controlled Trial. Diabetes Care 2018 Dec;41(12):2453-2462. [doi: 10.2337/dc18-0917] [Medline: 30305343]
49. Tavares J, Oliveira T. New Integrated Model Approach to Understand the Factors That Drive Electronic Health Record Portal Adoption: Cross-Sectional National Survey. J Med Internet Res 2018 Nov 19;20(11):e11032 [FREE Full text] [doi: 10.2196/11032] [Medline: 30455169]
50. Adjei Boakye E, Varble A, Rojek R, Peavler O, Trainer AK, Osazuwa-Peters N, et al. Sociodemographic Factors Associated With Engagement in Diabetes Self-management Education Among People With Diabetes in the United States. Public Health Rep 2018 Nov;133(6):685-691 [FREE Full text] [doi: 10.1177/0033354918794935] [Medline: 30223759]

51. Xie Z, Nacioglu A, Or C. Prevalence, Demographic Correlates, and Perceived Impacts of Mobile Health App Use Amongst Chinese Adults: Cross-Sectional Survey Study. JMIR Mhealth Uhealth 2018 Dec 26;6(4):e103 [FREE Full text] [doi: 10.2196/mhealth.9002] [Medline: 29699971]
52. Ernsting C, Dombrowski SU, Oedekoven M, Kanzler M, Kuhlmey A, Gellert P. Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey. J Med Internet Res 2017 Apr 05;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
53. Zhang Y, Liu C, Luo S, Xie Y, Liu F, Li X, et al. Factors Influencing Patients' Intentions to Use Diabetes Management Apps Based on an Extended Unified Theory of Acceptance and Use of Technology Model: Web-Based Survey. J Med Internet Res 2019 Aug 13;21(8):e15023 [FREE Full text] [doi: 10.2196/15023] [Medline: 31411146]
54. Deng Z, Hong Z, Ren C, Zhang W, Xiang F. What Predicts Patients' Adoption Intention Toward mHealth Services in China: Empirical Study. JMIR Mhealth Uhealth 2018 Aug 29;6(8):e172 [FREE Full text] [doi: 10.2196/mhealth.9316] [Medline: 30158101]

Abbreviations

FBG: fasting blood glucose
LCCP: Lilly Connected Care Program
PBG: postprandial blood glucose
SMBG: self-monitoring of blood glucose

Edited by G Eysenbach; submitted 13.12.19; peer-reviewed by K Fitzner, P Durneva, JD Ni; comments to author 08.01.20; revised version received 20.01.20; accepted 27.01.20; published 06.03.20.

Please cite as:
Zhang Y, Liu C, Luo S, Huang J, Li X, Zhou Z
Effectiveness of Lilly Connected Care Program (LCCP) App-Based Diabetes Education for Patients With Type 2 Diabetes Treated With Insulin: Retrospective Real-World Study
JMIR Mhealth Uhealth 2020;8(3):e17455
URL: https://mhealth.jmir.org/2020/3/e17455
doi:10.2196/17455
PMID:32141838

Original Paper

A Mobile-Based Intervention for Glycemic Control in Patients With Type 2 Diabetes: Retrospective, Propensity Score-Matched Cohort Study

Jing Li[1*], PhD; Li Sun[2*], PhD; Yaogang Wang[3*], PhD; Lichuan Guo[1], MA; Daiqing Li[1], PhD; Chang Liu[4], BS; Ning Sun[5], MBA; Zheng Xu[5], MD; Shu Li[3], PhD; Yunwen Jiang[3], BS; Yuan Wang[3], PhD; Shunming Zhang[3], MD; Liming Chen[1], PhD

[1]National Health Commission Key Laboratory of Hormones and Development, Tianjin Key Laboratory of Metabolic Diseases, Chu Hsien-I Memorial Hospital & Tianjin Institute of Endocrinology, Tianjin Medical University, Tianjin, China

[2]School of Nursing, Tianjin Medical University, Tianjin, China

[3]School of Public Health, Tianjin Medical University, Tianjin, China

[4]Yiducloud Technologies Co, Ltd, Beijing, China

[5]Yu-Tang Department, iHealth Labs Inc, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Liming Chen, PhD
National Health Commission Key Laboratory of Hormones and Development, Tianjin Key Laboratory of Metabolic Diseases
Chu Hsien-I Memorial Hospital & Tianjin Institute of Endocrinology
Tianjin Medical University
No 6, Huanruibei Road, Beichen District
Tianjin, 300134
China
Phone: 86 13920979401
Email: xfx22081@vip.163.com

Abstract

Background: Mobile-based interventions appear to be promising in ameliorating huge burdens experienced by patients with type 2 diabetes. However, it is unclear how effective mobile-based interventions are in glycemic management of patients with type 2 diabetes based on real-world evidence.

Objective: This study aimed to evaluate the effectiveness of a mobile-based intervention on glycemic control in patients with type 2 diabetes based on real-world population data.

Methods: This retrospective, propensity score-matched cohort study analyzed longitudinal data from a clinical electronic health database. The study population included 37,913 patients with type 2 diabetes at cohort entry between October 1, 2016, and July 31, 2018. A total of 2400 patients were matched 1:1, using propensity score matching, into the usual care and mobile health (mHealth) groups. The primary outcomes of glycemic control included control rates of glycated hemoglobin (HbA_{1c}), fasting blood glucose (FBG), and postprandial 2-hour blood glucose (P2BG). Mean values and variation trends of difference with 95% CI were the secondary outcomes. The general linear model was used to calculate repeated-measures analyses of variance to examine the differences between the two groups. Subgroup and sensitivity analyses were performed.

Results: Of the 2400 patients included in the analysis, 1440 (60.00%) were male and the mean age was 52.24 years (SD 11.56). At baseline, the control rates of HbA_{1c}, FBG, and P2BG in the mHealth and usual care groups were 45.75% versus 47.00% (*P*=.57), 38.03% versus 32.76% (*P*=.07), and 47.32% versus 47.89% (*P*=.83), respectively. At the 3-, 6-, 9-, and 12-month follow-ups, the mHealth group reported higher control rates of HbA_{1c} than did the usual care group: 69.97% versus 46.06% (*P*<.001), 71.89% versus 61.24% (*P*=.004), 75.38% versus 53.44% (*P*<.001), and 72.31% versus 46.70% (*P*<.001), respectively. At the four follow-up sessions, the control rates of FBG in the mHealth and usual care groups were statistically different: 59.24% versus 34.21% (*P*<.001), 56.61% versus 35.14% (*P*<.001), 59.54% versus 34.99% (*P*<.001), and 59.77% versus 32.83% (*P*<.001), respectively. At the four follow-up sessions, the control rates of P2BG in the mHealth group were statistically higher than in the usual care group: 79.72% versus 48.75% (*P*<.001), 80.20% versus 57.45% (*P*<.001), 81.97% versus 54.07% (*P*<.001), and

76.19% versus 54.21% (*P*=.001), respectively. At the four follow-up sessions, the percentages of HbA_{1c} reduction in the mHealth group were 8.66% (95% CI 6.69-10.63), 10.60% (95% CI 8.66-12.54), 10.64% (95% CI 8.70-12.58), and 8.11% (95% CI 6.08-10.14), respectively. At the four follow-up sessions, the percentages of P2BG reduction in the mHealth group were 8.44% (95% CI 7.41-10.73), 17.77% (95% CI 14.98-20.23), 16.23% (95% CI 13.05-19.35), and 16.91% (95% CI 13.17-19.84), respectively. Starting from the sixth month, the mean HbA_{1c} and P2BG values in the two groups increased slightly.

Conclusions: This mobile-based intervention delivered by a multidisciplinary team can better improve glycemic control rates of patients with type 2 diabetes than usual care. These effects were best sustained within the first 6 months. Starting from the sixth month, intensive management needs to be conducted to maintain long-term effectiveness of the mobile-based intervention.

(JMIR Mhealth Uhealth 2020;8(3):e15390) doi:10.2196/15390

KEYWORDS

mobile health; glycemic control; type 2 diabetes; propensity score matching

Introduction

The number of adults with diabetes worldwide increased from 108 million to 422 million between 1980 and 2014 [1], with a projected increase to 642 million by 2040 [2]. In China, the overall prevalence of diabetes in the adult population was estimated to be 11.6% in 2010 and was less than 1.0% in 1980 [3]. Patients with type 2 diabetes mellitus account for 90%-95% of those with diabetes [4]. Diabetes not only results in blindness, cardiovascular disease, kidney failure, and other long-term consequences that substantially impact quality of life and years of life lived with disability [5,6], but also increases the risk of cancer and all-cause mortality [7-11]. Therefore, the prevention and control of diabetes is becoming more and more important.

For people with diabetes, a series of cost-effective interventions can improve their health outcomes, regardless of what type of diabetes they may have [12-17]. These interventions mainly include glycemic control, combined with diet, physical activity, and, if necessary, medication; control of blood pressure and lipids to reduce cardiovascular risk and other complications; and regular screening for damage to the eyes, kidneys, and feet to facilitate early treatment [12,13]. Glycemic control through quarterly physician visits with measurements of glycated hemoglobin (HbA_{1c}), fasting blood glucose (FBG), and postprandial 2-hour blood glucose (P2BG) were recommended in professional treatment guidelines [14]. In addition, health education, counseling, self-management, and consistent follow-up were also important for people with diabetes [15-17]. Therefore, people with diabetes require access to systematic, ongoing, and organized care delivered by a multidisciplinary team of skilled health care providers.

Mobile health (mHealth), which is defined as the use of mobile and wireless technologies for health (ie, mobile phones or sensor technologies), aims to capitalize on the rapid uptake of information and communication technologies to improve health system efficiency and health outcomes [18-20]. This includes simple apps and complex technologies, including voice, text messaging (ie, short message service), multimedia message service, Bluetooth technology, and others [21]. These advances, combined with changing patient attitudes toward self-testing, as well as an increased interest in wearable biosensors, are enticing health care providers to shift toward the paradigm of *P4* medicine: predictive, pre-emptive, personalized, and participatory [22]. The characteristics of mobility, instantaneous access, and direct communication of mHealth allow for faster transfer of health information, which in turn supports patient management. mHealth is a promising tool for delivering interventions designed to promote lifestyle management of patients with type 2 diabetes. Use of mHealth often includes the possibility of sharing data between health professionals and their patients with diabetes, which could enhance the support to improve their self-management [23,24]. Successful use of mHealth technology requires an active user and cooperation among health professionals [25]. That is, the technology's effectiveness is often determined by the way in which it is provided to patients or practitioners, how it is supported or taught, and how mHealth technology is added to clinical work or daily life [26].

Previous studies have shown that mobile-based interventions developed for diabetes holistic management have some effects [23-25,27,28]. A systematic review included a meta-analysis of 14 randomized trials aiming to investigate the effect of apps on HbA_{1c} in the self-management of diabetes; these studies showed that the mean reduction in HbA_{1c} among participants using an app compared with control group participants was 0.49% (95% CI 0.30-0.68; I^2=10%) [23]. Another systematic review of high-quality review articles and meta-analysis, which focused on utilizing technology in diabetes self-management education and support services, found that technology-enabled diabetes self-management solutions significantly improved HbA_{1c} and four key elements emerged as essential for improved HbA_{1c}: (1) communication, (2) patient-generated health data, (3) education, and (4) feedback [24]. Although the majority of these interventions showed improvement on primary endpoints [25,27,28], whether results will drive substantial clinical adoption is unknown because small studies, even if randomized, are unlikely to be significantly powered to demonstrate meaningful real-world effects [20]. Therefore, real-world evidence and performance data of mobile-based interventions are needed to demonstrate value or motivate stakeholder adoption.

Based on real-world population data from a clinical electronic health database, this study aimed to evaluate the effectiveness of a mobile-based intervention on glycemic control in patients with type 2 diabetes and to explore the change in trends of glycemic parameters in the short and long term.

Methods

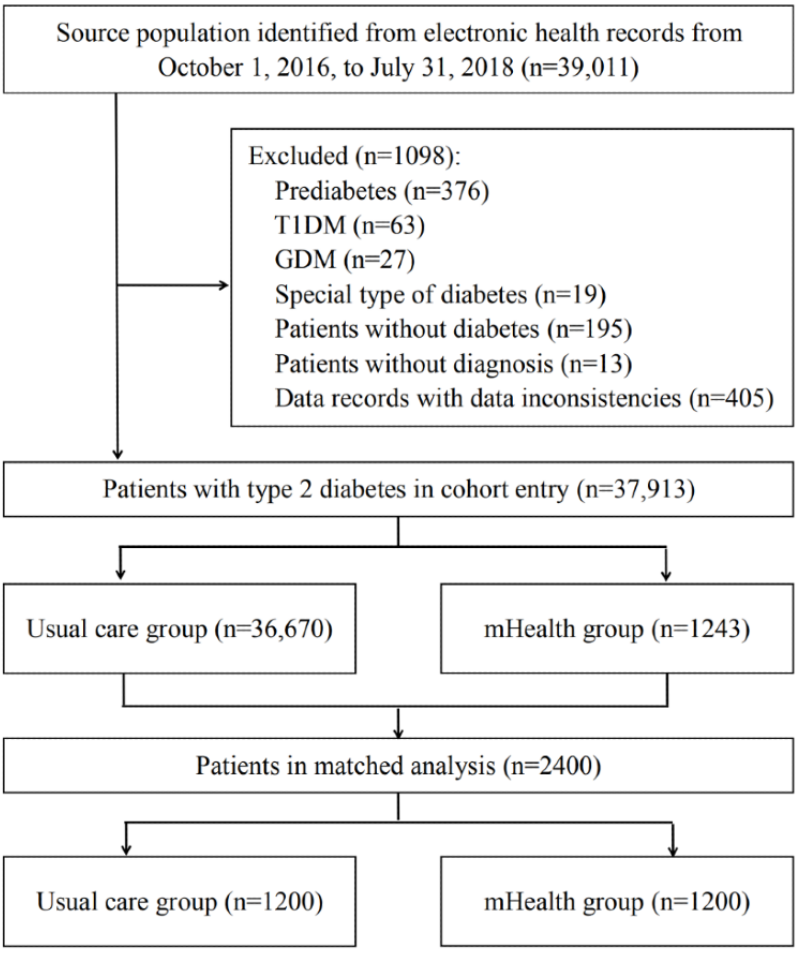
Study Design

We conducted a retrospective, propensity score-matched cohort study using electronic health data from a clinical database in Tianjin, China. This clinical database was established in June 2014. The database contained longitudinal outpatient records of patients in five primary care practices and one tertiary care hospital specializing in diabetes, including their demographics, primary and secondary diagnoses, and clinical examination results.

Cohort Selection

The study cohort included 37,913 patients with type 2 diabetes who were registered in this clinical database between October 1, 2016, and July 31, 2018. The flowchart for cohort selection is shown in Figure 1.

Figure 1. Flowchart for cohort selection. GDM: gestational diabetes mellitus; T1DM: type 1 diabetes mellitus.



All of the patients who were registered in the clinical database between October 1, 2016, and July 31, 2018, were included in our source population. In the end, we identified 39,011 individuals with 1,793,841 records from the clinical database. The unique ID numbers were used to identify the records of the patients with different outpatient numbers in different clinical settings. Using Microsoft Visual Basic 6.0, we developed sophisticated applications to extract and filter these data. After source population selection, the following exclusion criteria were applied: (1) prediabetes, (2) type 1 diabetes, (3) gestational diabetes, (4) special type of diabetes, (5) patients without diabetes, (6) patients without diagnosis, and (7) data records with data inconsistencies. Finally, we identified 37,913 patients with type 2 diabetes.

This unmatched cohort was divided into two groups, including the usual care group (n=36,670) and the mHealth group (n=1243). This was an observational study originating from the real world with no constraints on the cohort entry of participants in either group.

Interventions

The usual care group received standard medical care for patients with type 2 diabetes. Every 3 months, patients in this group underwent regular reviews to re-examine HbA_{1c}, FBG, and P2BG levels. These lab examination results were considered as evidence to support doctors' decisions to adjust medications. Meanwhile, patients in this group received routine health education at each session.

In addition to usual care, the mHealth group received a mobile-based intervention, which was continuous, real-time, personalized health care delivered by a multidisciplinary team consisting of doctors, nurses, health educators, and dietitians.

This mobile-based intervention was based on a unified diabetes care system, which consisted of a mobile app, smart wearable medical devices (eg, wireless glucose monitor, wireless blood pressure monitor, pulse oximeter, and body composition scale), a Web platform, and a data-sharing cloud platform. Patients with type 2 diabetes in the mHealth group followed new flows in the clinical settings (see Multimedia Appendix 1). Patients' management and education in the mHealth group extended from the clinic to home. They used the wireless glucose monitors and app to perform glucose checks at home. When they experienced hypoglycemia or hyperglycemia, the app could provide tips to help them regulate their glucose levels. Patients also sent their results immediately to the support team about what to do. The care team and the service support team members would be notified when the patient was experiencing abnormal glucose levels. They then phoned the patient to inquire about their recent medication, diet, and exercise, and to help the patient in analyzing possible reasons for the abnormal glucose level. If necessary, they would invite the patient for further in-clinic consultation or guide the patient to adjust their diet or exercise by phone. Patients could also record their meals in the app and get feedback from the service support team. According to an image or a description of the food, the team would provide an overall rating of the meal, comments on portion and nutrition, and suggestions on how to do better. Patients could log their exercise type and duration into the app. The service support team created updated knowledge covering blood glucose, blood pressure, food, fitness, oral medication, insulin, psychology, and complications in the form of articles, videos, and attractive posters. Patients had access to this educational information whenever and wherever possible.

Outcome Definition

The primary outcomes were control rates of HbA_{1c}, FBG, and P2BG at baseline and at 3-, 6-, 9-, and 12-month follow-ups. We identified the control rates according to guidelines for the prevention and control of type 2 diabetes in China [29]. Control objectives were defined as follows: (1) HbA_{1c} <7%, (2) 4.4 mmol/L< FBG <7.0 mmol/L, and (3) P2BG <10 mmol/L. We also considered mean values and variation trends of difference (VTD) with 95% CI, separately, as secondary outcomes. The formula for calculating VTD was as follows:

$$VTD = (Value_n - Value_{baseline})/Value_{baseline} \times 100\%$$

where $Value_n$ and $Value_{baseline}$ denoted the sample values of HbA_{1c}, FBG, and P2BG at *n*-month (n=3, 6, 9, and 12) follow-up and baseline, respectively [30]. If VTD was positive, it represented the percentage of increase; if VTD was negative, it represented the percentage of reduction.

Covariates

Demographic and chronic disease covariates included sex, age, comorbidity (ie, hyperlipidemia and hypertension), P2BG, FBG, HbA_{1c}, and low-density lipoprotein (LDL) cholesterol. App use-related covariates included times of FBG and P2BG self-testing, diet records, exercise records, and out-of-hospital follow-up.

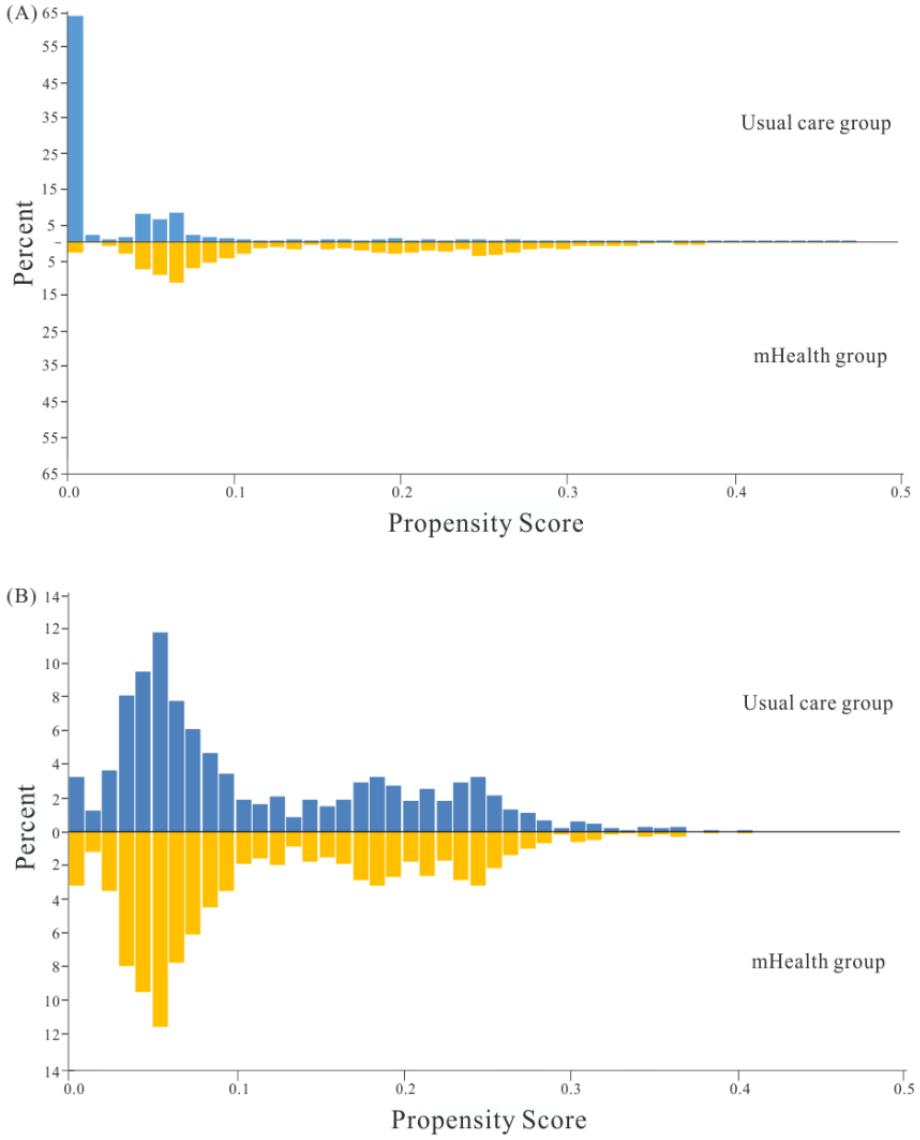
Statistical Analysis

To control for the nonrandom assignment of patients, a logistic regression model that predicted the likelihood of being included in the mHealth group was constructed and used as the propensity score. Patients were then matched 1:1, using propensity score matching, into the usual care group and mHealth group. We selected all of the common available variables for two-group matching [31], including sex, age group, comorbidity (ie, hyperlipidemia and hypertension), HbA_{1c} level, and LDL cholesterol level. The propensity score-matching tolerance was 0.005. No replacement was allowed, and patients were matched only once. Standardized differences with mirror histograms before and after matching are shown in Figure 2 and Multimedia Appendix 2. We evaluated the balances of matched covariates with standardized differences [29] and considered differences of less than 10% to be matched sufficiently [32,33].

We presented categorical variables as numbers (percentages) and continuous variables as means (SDs) or 95% CIs, or as medians (IQRs), as appropriate. Descriptive statistics were used to analyze the patient demographics. Binary or categorical outcome measures were analyzed using the chi-square test and continuous measures were analyzed using the *t* test or a nonparametric equivalent (eg, Wilcoxon rank test). We used the general linear model to calculate repeated-measures analyses of variance to examine mean differences of two groups at baseline and at 3-, 6-, 9-, and 12-month follow-ups. Subgroup analyses to explore the effects of this mobile-based intervention in different patient subgroups were undertaken for the primary and secondary outcomes. The subgroups, specified in the statistical analysis, included patient demographics: sex, age group, hyperlipidemia, and hypertension. The total proportion of missing values at the 12-month follow-up was 12.3%. The proportions of missing data at each data point were 4.5% (3-month follow-up), 7.1% (6-month follow-up), and 9.5% (9-month follow-up). Expectation maximization was used to estimate the missing values of continuous variables. A sensitivity analysis was performed by repeating our primary analysis but excluding patients with hyperlipidemia or hypertension.

We determined statistical significance using a two-tailed *P* value of less than .05. All of the statistical analyses were carried out using SPSS Statistics for Windows, version 25.0 (IBM Corp).

Figure 2. Mirror histograms. (A) Before match. (B) After match.



Results

Patient Demographics

Of the 39,011 patients, 37,913 met the selection criteria for additional analysis (see Figure 1). In the unmatched cohort, the proportion of male patients was 53.41% (20,248/37,913), patients' mean age was 57.94 years (SD 12.10), and 88.13% of patients were 36-74 years of age. The proportion of patients with hyperlipidemia was 20.52% (7779/37,913), and the proportion of patients with hypertension was 8.11% (3073/37,913). The mean HbA_{1c} level was 7.86% (SD 1.25) and the mean LDL cholesterol level was 3.37 mmol/L (SD 0.65). Table 1 shows the baseline demographics of patients with type 2 diabetes in unmatched and propensity score-matched cohorts. There were significant differences in demographics or glycemic parameters between the usual care group and the mHealth group.

A propensity score match was then performed and 2400 patients were matched 1:1. After matching, covariates were well balanced and we did not observe any significant differences between groups (see Table 1). In the propensity score-matched cohort, the proportion of male patients was 60.00% (1440/2400), the mean age of patients was 52.24 years (SD 11.56), and 90.13% of patients were 36-74 years of age. A total of 48.54% (1165/2400) of patients had a comorbidity of hyperlipidemia, and 44.38% (1065/2400) of patients had a comorbidity of hypertension. The mean HbA_{1c} level was 7.76% (SD 1.39) and the mean LDL cholesterol level was 3.27 mmol/L (SD 0.79).

Table 1. Baseline demographics of patients with type 2 diabetes in unmatched and propensity score-matched cohorts.

Characteristic	Unmatched cohort				Propensity score-matched cohort			
	mHealth group (n=1243)	Usual care group (n=36,670)	*P* value	Std diff[a]	mHealth group (n=1200)	Usual care group (n=1200)	*P* value	Std diff
Sex, n (%)								
Male	777 (62.51)	19,471 (53.10)	<.001	19.13	738 (61.50)	702 (58.50)	.15	3.80
Female	466 (37.49)	17,199 (46.90)			462 (38.50)	498 (41.50)		
Age group (years), n (%)								
≤35	80 (6.44)	2086 (5.69)	<.001	29.17	73 (6.08)	96 (8.00)	.07	7.31
36-59	657 (52.85)	15,914 (43.40)		19.10	631 (52.58)	634 (52.83)		0.40
60-74	464 (37.33)	16,379 (44.66)		15.10	454 (37.84)	444 (37.00)		1.67
≥75	42 (3.38)	2291 (6.25)		13.13	42 (3.50)	26 (2.17)		7.78
Comorbidity, n (%)								
Hyperlipidemia	647 (52.05)	7132 (19.45)	<.001	72.67	604 (50.33)	561 (46.75)	.09	7.00
Hypertension	593 (47.71)	2480 (6.76)	<.001	102.25	550 (45.83)	515 (42.92)	.16	5.80
Biochemical indicator, mean (SD)								
HbA_{1c}[b] level (%)	7.82 (1.60)	7.86 (1.24)	<.001	3.16	7.83 (1.60)	7.70 (1.15)	.26	9.35
LDL[c] cholesterol (mmol/L)	3.26 (0.88)	3.38 (0.64)	<.001	15.10	3.29 (0.88)	3.25 (0.69)	.71	5.06

[a]Std diff: standardized difference.

[b]HbA_{1c}: glycated hemoglobin.

[c]LDL: low-density lipoprotein.

Until July 31, 2018, the total number of times of starting up the app, self-monitoring of glycemic parameters, diet recording, exercise recording, and out-of-hospital follow-ups were 80,129; 172,355; 17,860; 4464; and 5264, respectively; the median follow-up time was 457 days.

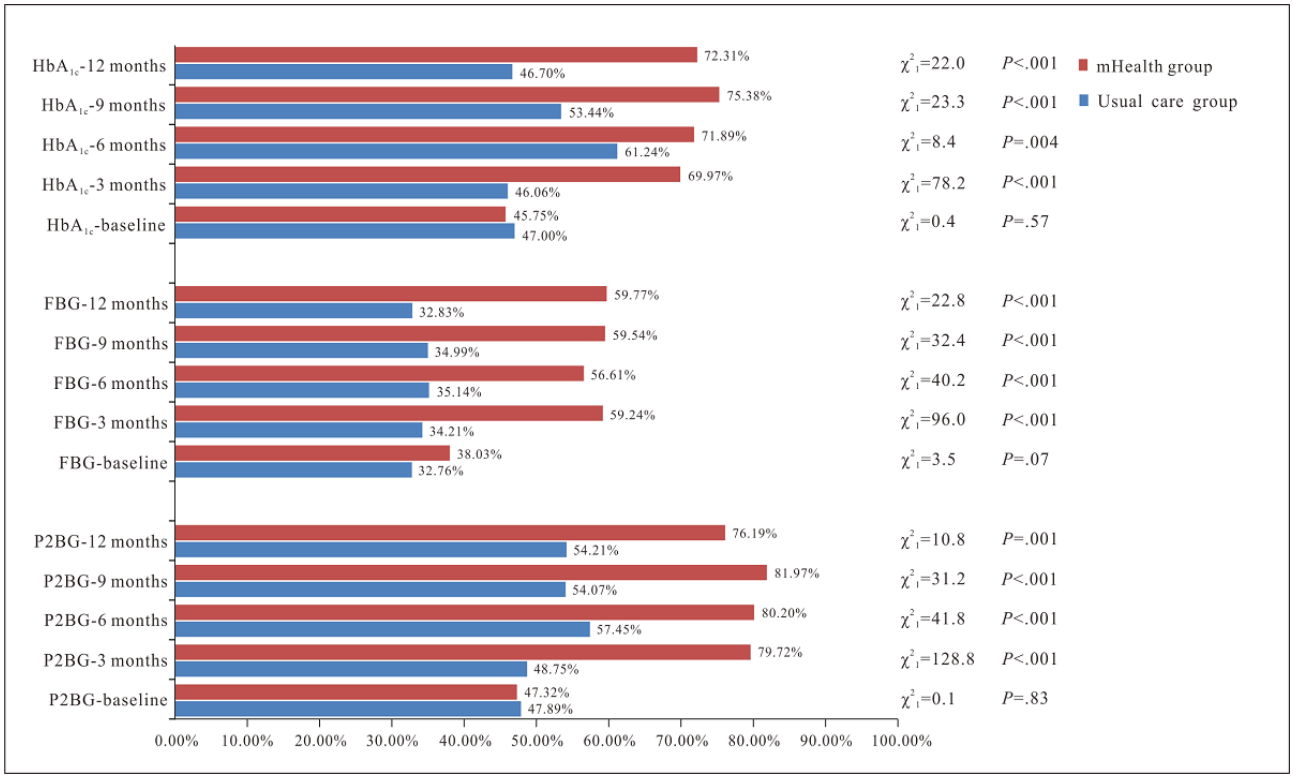
Control Rates of Glycemic Parameters

At baseline, the control rates of HbA_{1c}, FBG, and P2BG in the mHealth and usual care groups were 45.75% versus 47.00% (*P*=.57), 38.03% versus 32.76% (*P*=.07), and 47.32% versus 47.89% (*P*=.83), respectively. The control rates of HbA_{1c}, FBG, and P2BG in both groups at different follow-up sessions are shown in Figure 3.

At the 3-, 6-, 9-, and 12-month follow-ups, the mHealth group reported higher control rates of HbA_{1c} than usual care, which were 69.97% versus 46.06% (*P*<.001), 71.89% versus 61.24% (*P*=.004), 75.38% versus 53.44% (*P*<.001), and 72.31% versus 46.70% (*P*<.001), respectively. Differences in the control rates between the two groups at these four follow-up sessions were 23.91%, 10.65%, 21.94%, and 25.61%, respectively. At the 9-month follow-up, the mHealth group reported the highest control rate of HbA_{1c}, which was 75.38%.

Additionally, at the 3-, 6-, 9-, and 12-month follow-ups, the control rates of FBG in the mHealth and usual care groups were statistically different, which were 59.24% versus 34.21% (*P*<.001), 56.61% versus 35.14% (*P*<.001), 59.54% versus 34.99% (*P*<.001), and 59.77% versus 32.83% (*P*<.001), respectively. The control rates of P2BG in the mHealth group were statistically higher than in the usual care group, which were 79.72% versus 48.75% (*P*<.001), 80.20% versus 57.45% (*P*<.001), 81.97% versus 54.07% (*P*<.001), and 76.19% versus 54.21% (*P*=.001), respectively.

Figure 3. Control rates of glycated hemoglobin (HbA_{1c}), fasting blood glucose (FBG), and postprandial 2-hour blood glucose (P2BG) in the mHealth and usual care groups at different follow-up sessions.



Mean Values of Glycemic Parameters

Table 2 shows the effects of this mobile-based intervention on glycemic parameters. The mean values of HbA_{1c}, FBG, and P2BG in the mHealth group were significantly lower than those in the usual care group at the 3-, 6-, 9-, and 12-month follow-ups (P<.01). The P values of the *month* factor were less than .001, which meant that the HbA_{1c}, FBG, and P2BG levels changed with time. The *group* and *month* factors had interaction effects in the mean values of HbA_{1c}, FBG, and P2BG (P<.01), which meant that the effect of the *time* factor varied with the *group*. These results identified improved effects due to this mobile-based intervention on changes of the HbA_{1c}, FBG, and P2BG mean values.

Multimedia Appendices 3-7 show that, compared with usual care, at the 3-, 6-, 9-, and 12-month follow-ups, both sexes in the mHealth group reported significantly lower mean values of HbA_{1c}, FBG, and P2BG (P<.05), although we did not observe any significant differences in the P2BG mean values of female participants between the two groups at the 12-month follow-up (P=.20). Patients aged 36-74 years in the mHealth group had steadily lower HbA_{1c}, FBG, and P2BG mean values than those in the usual care group (P<.05). No statistically significant difference was observed in P2BG mean values of patients aged 36-59 years between the two groups at the 12-month follow-up (P=.09). Patients younger than 35 or older than 75 years of age in the mHealth group reported unstable variation trends of mean values.

Table 2. Effects of the mobile-based intervention on glycemic parameters.

Variables and group	Measurement session					F months	*P* value	F groups	*P* value	F months × group	*P* value
	Baseline	3 months	6 months	9 months	12 months						
HbA_{1c}[a] (%), mean (SD)											
mHealth group	7.83 (1.60)	6.70 (0.73)	6.60 (0.61)	6.44 (0.59)	6.75 (0.76)	11.822	<.001	0.058	.08	5.905	.003
Usual care group	7.70 (1.15)	7.36 (1.17)	6.82 (0.64)	6.97 (0.63)	7.12 (0.64)						
Z	-1.123	-20.382	-18.592	-19.922	-9.657						
P value	.26	<.001	<.001	<.001	<.001						
FBG[b] (mmol/L), mean (SD)											
mHealth group	8.34 (2.41)	6.51 (1.27)	6.74 (2.11)	6.68 (1.49)	6.89 (1.52)	9.614	<.001	16.425	<.001	5.762	<.001
Usual care group	8.68 (2.34)	8.53 (2.37)	8.45 (2.45)	8.38 (2.39)	8.47 (2.68)						
Z	-1.326	-15.268	-17.315	-14.831	-5.755						
P value	.20	<.001	<.001	<.001	<.001						
P2BG[c] (mmol/L), mean (SD)											
mHealth group	11.14 (4.40)	8.03 (1.75)	7.75 (1.90)	7.89 (1.77)	8.29 (2.38)	12.424	<.001	9.566	.002	7.193	<.001
Usual care group	10.36 (2.94)	9.99 (2.99)	9.85 (2.56)	9.93 (2.83)	9.97 (2.91)						
Z	-1.575	-23.995	-20.921	-16.243	-3.149						
P value	.12	<.001	<.001	<.001	.002						

[a]HbA_{1c}: glycated hemoglobin.

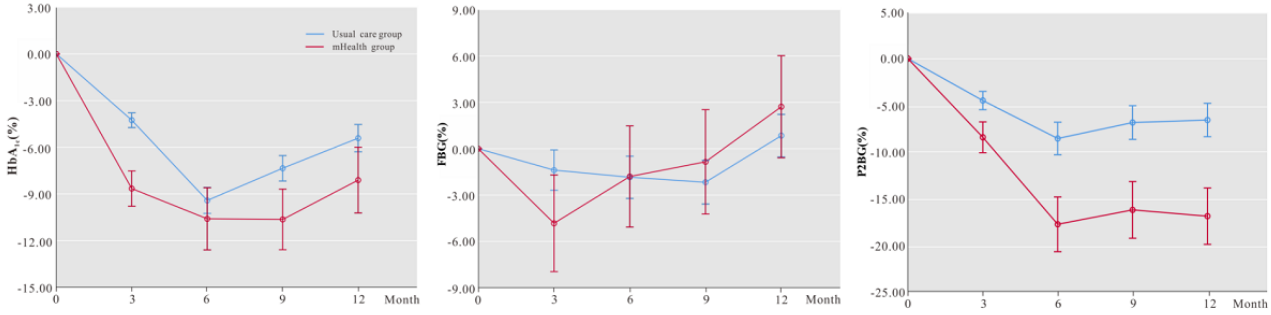
[b]FBG: fasting blood glucose.

[c]P2BG: postprandial 2-hour blood glucose.

Variation Trends of Difference for Glycemic Parameters

Figure 4 shows the variation trends of difference for glycemic parameters between the two groups. At the 3-, 6-, 9-, and 12-month follow-ups, the percentages of HbA_{1c} reduction in mHealth group were 8.66% (95% CI 6.69-10.63), 10.60% (95% CI 8.66-12.54), 10.64% (95% CI 8.70-12.58), and 8.11% (95% CI 6.08-10.14), respectively; the percentages of P2BG reduction in the mHealth group were 8.44% (95% CI 7.41-10.73), 17.77% (95% CI 14.98-20.23), 16.23% (95% CI 13.05-19.35), and 16.91% (95% CI 13.17-19.84), respectively. Equally important was that, after 6 months, the declines in HbA_{1c} and P2BG of the two groups decreased, whereas the mHealth group experienced larger decreases in HbA_{1c} and P2BG than the usual care group. At the 3-month follow-up, the reduction of FBG in the mHealth group was larger than in the usual care group (4.83% vs 1.38%). However, starting from the sixth month, the reductions of FBG in the usual care group were larger than in the mHealth group.

Figure 4. Variation trends of glycated hemoglobin (HbA_{1c}), fasting blood glucose (FBG), and postprandial 2-hour blood glucose (P2BG) differences.



Sensitivity Analysis

We conducted sensitivity analysis, where we excluded 1587 patients with hyperlipidemia or hypertension. In the mHealth group at the baseline and the four follow-up sessions, the control rates of HbA_{1c} were 36.57%, 72.77%, 76.92%, 80.00%, and 75.00%, respectively; the control rates of FBG were 44.44%, 65.79%, 65.74%, 59.38%, and 65.63%, respectively; and the control rates of P2BG were 54.02%, 82.63%, 77.45%, 83.33%, and 72.41%, respectively. There were no significant differences between the sensitivity analysis and the primary analysis (χ^2=10.0 *P*=.35), so we presented only the results of the primary analysis.

Discussion

Principal Findings

In this study, a total of 2400 patients with type 2 diabetes were matched 1:1, using propensity score matching, into the usual care and mHealth groups. A total of 60% of the patients were male and more than half were 36-59 years of age. These demographics were similar to the population in previous studies [29,34]. Our results showed the improvement in control rates of HbA_{1c}, FBG, and P2BG for patients with type 2 diabetes in the mHealth group compared with those in the usual care group. These effects were best sustained within the first 6 months. Starting from the sixth month, the mean HbA_{1c} and P2BG values in the two groups increased slightly.

Comparison With Prior Work

The role of glycemic control in preventing the development and progression of complications has been proven in diabetes [35-38], with an especially strong relationship identified between intensive glycemic control and diabetic complications and mortality. In general, a target HbA_{1c} level of less than 7% is optimal, according to diabetes guidelines [14]. Each 1% of mean HbA_{1c} value reduction has been associated with a 21% reduction in the risk of diabetes-related complications [39]. A recent study on the legacy effect of early glycemic control on future complications in type 2 diabetes showed that, compared with an HbA_{1c} of less than 6.5% for the 0-1-year early exposure period, HbA_{1c} levels of 6.5% or higher were associated with increased microvascular and macrovascular events, and HbA_{1c} levels of 7.0% or higher were associated with increased mortality [40]. However, a recent meta-analysis demonstrated that HbA_{1c} target achievement is low, with a pooled average of 43% worldwide [41], both in primary and secondary care settings. In 2013, among Chinese patients with diabetes, only 39.7% of those treated had adequate glycemic control [34]. The reason for this low target achievement, despite the expanding arsenal of glucose-lowering interventions, remains unclear [42]. Our study found that the control rates of HbA_{1c}, FBG, and P2BG in the mHealth group were higher than those in the usual care group, which were much higher than the average level worldwide [41]. These findings confirmed the effectiveness of this mobile-based intervention on glycemic control in patients with type 2 diabetes. Our study also found that at different follow-up sessions, both sexes in the mHealth group reported significantly lower mean values of HbA_{1c}, FBG, and P2BG. Patients aged 36-74 years in the mHealth group had steadily lower HbA_{1c}, FBG, and P2BG mean values than those in the usual care group, and patients younger than 35 or older than 75 years old in the mHealth group reported unstable variation trends of mean values. There may be many reasons for poor glycemic control in patients older than 75 years of age in the mHealth group, including decreased self-management ability, inadequate exercise, irregular glycemic monitoring, and poor convenience in using apps, among other reasons [21,23,43]. For patients younger than 35 years old, the main reason may be poor compliance of patients and insufficient understanding of the importance of glycemic monitoring [24,25]. Especially for young patients, poor parental health literacy is the main reason [44,45].

However, some studies have found that even if blood glucose is effectively controlled, the occurrence and development of complications cannot be improved or reversed [46-48]. Researchers believe that this is due to the “metabolic memory” effect of hyperglycemia [46-48]. “Metabolic memory” effect refers to the persistent damage of early hyperglycemia to tissues and organs of diabetic patients, even though the glycemic control is good [48]. A growing body of experimental evidence supports the concept that the risk for diabetes complications may be linked to oxidative stress, nonenzymatic glycosylation of proteins, epigenetic changes, and chronic inflammation, laying the foundation for the “metabolic memory” theory [46]. From a clinical standpoint, the “metabolic memory” theory supports the need for very early aggressive treatment, with the goal of normalizing metabolic control as soon as possible, especially blood glucose. Therefore, achieving glycemic control targets as soon as possible and maintaining glycemic control for a long time have significantly positive effects in the prevention of complications [14]. The treatment strategy of diabetes should

be changed from strict glycemic control to strict glycemic control at the early stage. The mobile-based intervention in this study seems to offer a promising option to implement this strategy. One meta-analysis of 35 randomized controlled trials found that an internet-based or mobile-based intervention duration of 3 months or less yielded optimal performance [49]. Our study also found that the glycemic control rates of the mHealth group were higher than those of the usual care group at the 3-, 6-, 9-, and 12-month follow-ups. These findings were not only consistent with previous studies, but also illustrated that the mobile-based intervention had generated a statistically significant improvement on glycemic control in the short and long term [47,49]. Therefore, implementation of mobile-based interventions could be a promising strategy for glycemic control of patients with diabetes not only at the early stage, but also in the long term.

In our study, the mobile-based intervention was designed to provide continuous, real-time, personalized health care for patients with diabetes, and it was delivered by a multidisciplinary team consisting of doctors, nurses, health educators, and dietitians. With the help of mobile technologies, this intervention provides a solution for diabetes management that includes the following: (1) a simple and intuitive way of vital data collection, (2) automatic in-hospital exam data and at-home data consolidation, (3) convenient and timely communication with care team professionals, and (4) continuous and vivid diabetes education, both in person and through multimedia. Based on the hardware equipment and professional support team, we have realized real-time guidance and management for diabetic patients; meanwhile, we have also collected a large amount of sample data. These data from the real world reflected the effectiveness of this mobile-based intervention. Notably, we found that, starting from the sixth month, the glycemic control of patients with type 2 diabetes in the mHealth group began to fluctuate slightly. This is a reminder that intensive management needs to be conducted to maintain the long-term effectiveness of this mobile-based intervention from the sixth month.

Strengths and Limitations

A major strength of this study was the high-quality, continuously updated, clinical database of electronic medical records that provided a large sample size and reflected real-world clinical conditions. In addition, in our study, propensity score matching was used to control the confounding factors between the two groups. Propensity score matching could reduce the bias resulting from confounding variables; this approach attempted to mimic randomization by creating a sample of units that received the mobile-based intervention that is comparable, on all observed covariates, to a sample of units that received usual care.

This study had several limitations. First, as a result of its retrospective nature, we may not have addressed unobserved confounders in propensity score matching. Therefore, selection bias may exist in this research. For this reason, we used the propensity score matching to balance the common available covariates of the two groups, including sex, age, comorbidity (ie, hyperlipidemia and hypertension), HbA_{1c} level, and LDL cholesterol level. Second, there are inherent limitations as to what data are recorded in the clinical medical records. For instance, the clinical medical records of patients in the usual care group did not include some demographic information, such as education level, economic level, and occupation, among others, as well as anthropometry data, such as height, weight, systolic blood pressure, and diastolic blood pressure, among others. Cognitive function was not evaluated for either group. Third, in this study, the total proportion of missing values was 12.3%, and the missing values were on continuous variables, including HbA_{1c}, FBG, P2BG, and LDL cholesterol. In order to decrease the amount of bias in the data, we used expectation maximization to estimate the missing values of continuous variables.

Conclusions

This mobile-based intervention delivered by a multidisciplinary team to promote glycemic control of patients with type 2 diabetes led to increases in the control rates of HbA_{1c}, FBG, and P2BG. These effects were best sustained within the first 6 months. It is noteworthy that, starting from the sixth month, intensive management might need to be conducted to maintain long-term effectiveness of this mobile-based intervention.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (#91746205 and #71673199), the Tianjin Science and Technology Project (#15ZXHLSY00460 and #18ZXZNSY00280), the Natural Science Foundation of Tianjin City (#18JCYBJC26100), and key social science projects of the Tianjin Education Commission (#2019JWZD54).

Authors' Contributions

LC and YW (Yaogang Wang) contributed to the conception and design of the study. JL, LG, DL, CL, NS, and ZX acquired the data. JS, LS, SL, YJ, YW (Yuan Wang), and SZ analyzed the data and all authors interpreted the data. JL, LS, YW (Yaogang Wang), and LC drafted the manuscript and all authors were involved in critical revision and approval of the final manuscript. YW (Yaogang Wang) and LC contributed equally to this paper as corresponding authors. (Yaogang Wang, PhD, School of Public Health, Tianjin Medical University No 22, Qixiangtai Road, Heping District Tianjin, 300070, China, Phone: 86 13820046130, Email: wyg@tmu.edu.cn) The corresponding authors attest that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1
In-clinic user flow of the mHealth group.
[DOC File , 4589 KB - mhealth_v8i3e15390_app1.doc]

Multimedia Appendix 2
Standardized differences before and after propensity score matching.
[DOC File , 166 KB - mhealth_v8i3e15390_app2.doc]

Multimedia Appendix 3
Variation trends of glycated hemoglobin (HbA1c), fasting blood glucose (FBG), and postprandial 2-hour blood glucose (P2BG) mean values (sex groups).
[DOC File , 532 KB - mhealth_v8i3e15390_app3.doc]

Multimedia Appendix 4
Variation trends of glycated hemoglobin (HbA1c), fasting blood glucose (FBG), and postprandial 2-hour blood glucose (P2BG) mean values (age groups).
[DOC File , 929 KB - mhealth_v8i3e15390_app4.doc]

Multimedia Appendix 5
Subgroup analysis of glycated hemoglobin (HbA1c) (%) between usual care and mHealth groups.
[DOC File , 71 KB - mhealth_v8i3e15390_app5.doc]

Multimedia Appendix 6
Subgroup analysis of fasting blood glucose (FBG) (mmol/L) between usual care and mHealth groups.
[DOC File , 69 KB - mhealth_v8i3e15390_app6.doc]

Multimedia Appendix 7
Subgroup analysis of postprandial 2-hour blood glucose (P2BG) (mmol/L) between usual care and mHealth groups.
[DOC File , 69 KB - mhealth_v8i3e15390_app7.doc]

References

1. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in diabetes since 1980: A pooled analysis of 751 population-based studies with 4.4 million participants. Lancet 2016 Apr 09;387(10027):1513-1530 [FREE Full text] [doi: 10.1016/S0140-6736(16)00618-8] [Medline: 27061677]
2. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract 2017 Jun;128:40-50. [doi: 10.1016/j.diabres.2017.03.024] [Medline: 28437734]
3. National Diabetes Research Group. A mass survey of diabetes mellitus in a population of 300,000 in 14 provinces and municipalities in China (author's transl) [Article in Chinese]. Zhonghua Nei Ke Za Zhi 1981 Nov;20(11):678-683. [Medline: 7341098]
4. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2014 Jan;37 Suppl 1:S81-S90. [doi: 10.2337/dc14-S081] [Medline: 24357215]
5. The Lancet. Type 2 diabetes: The urgent need to protect young people. Lancet 2018 Dec 01;392(10162):2325. [doi: 10.1016/S0140-6736(18)33015-0] [Medline: 30527598]
6. Xu G, Liu B, Sun Y, Du Y, Snetselaar LG, Hu FB, et al. Prevalence of diagnosed type 1 and type 2 diabetes among US adults in 2016 and 2017: Population based study. BMJ 2018 Sep 04;362:k1497 [FREE Full text] [doi: 10.1136/bmj.k1497] [Medline: 30181166]
7. Harding JL, Shaw JE, Peeters A, Cartensen B, Magliano DJ. Cancer risk among people with type 1 and type 2 diabetes: Disentangling true associations, detection bias, and reverse causation. Diabetes Care 2015 Feb;38(2):264-270. [doi: 10.2337/dc14-1996] [Medline: 25488912]
8. Preis SR, Hwang S, Coady S, Pencina MJ, D'Agostino RB, Savage PJ, et al. Trends in all-cause and cardiovascular disease mortality among women and men with and without diabetes mellitus in the Framingham Heart Study, 1950 to 2005. Circulation 2009 Apr 07;119(13):1728-1735 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.108.829176] [Medline: 19307472]

9. Baena-Díez JM, Peñafiel J, Subirana I, Ramos R, Elosua R, Marín-Ibañez A, FRESCO Investigators. Risk of cause-specific death in individuals with diabetes: A competing risks analysis. Diabetes Care 2016 Nov;39(11):1987-1995. [doi: 10.2337/dc16-0614] [Medline: 27493134]
10. Gregg EW, Li Y, Wang J, Burrows NR, Ali MK, Rolka D, et al. Changes in diabetes-related complications in the United States, 1990-2010. N Engl J Med 2014 Apr 17;370(16):1514-1523. [doi: 10.1056/NEJMoa1310799] [Medline: 24738668]
11. Tsilidis KK, Kasimis JC, Lopez DS, Ntzani EE, Ioannidis JP. Type 2 diabetes and cancer: Umbrella review of meta-analyses of observational studies. BMJ 2015 Jan 02;350:g7607 [FREE Full text] [doi: 10.1136/bmj.g7607] [Medline: 25555821]
12. Gaede P, Lund-Andersen H, Parving H, Pedersen O. Effect of a multifactorial intervention on mortality in type 2 diabetes. N Engl J Med 2008 Feb 07;358(6):580-591. [doi: 10.1056/NEJMoa0706245] [Medline: 18256393]
13. Wei M, Gibbons LW, Kampert JB, Nichaman MZ, Blair SN. Low cardiorespiratory fitness and physical inactivity as predictors of mortality in men with type 2 diabetes. Ann Intern Med 2000 Apr 18;132(8):605-611. [doi: 10.7326/0003-4819-132-8-200004180-00002] [Medline: 10766678]
14. Garber AJ, Abrahamson MJ, Barzilay JI, Blonde L, Bloomgarden ZT, Bush MA, et al. Consensus Statement by the American Association of Clinical Endocrinologists and American College of Endocrinology on the comprehensive type 2 diabetes management algorithm: 2019 executive summary. Endocr Pract 2019 Jan;25(1):69-100. [doi: 10.4158/CS-2018-0535] [Medline: 30742570]
15. Dong Y, Wang P, Dai Z, Liu K, Jin Y, Li A, et al. Increased self-care activities and glycemic control rate in relation to health education via WeChat among diabetes patients: A randomized clinical trial. Medicine (Baltimore) 2018 Dec;97(50):e13632 [FREE Full text] [doi: 10.1097/MD.0000000000013632] [Medline: 30558051]
16. Bluml BM, Kolb LE, Lipman R. Evaluating the impact of year-long, augmented diabetes self-management support. Popul Health Manag 2019 Dec;22(6):522-528 [FREE Full text] [doi: 10.1089/pop.2018.0175] [Medline: 30668228]
17. Ayre J, Bonner C, Bramwell S, McClelland S, Jayaballa R, Maberly G, et al. Factors for supporting primary care physician engagement with patient apps for type 2 diabetes self-management that link to primary care: Interview study. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11885 [FREE Full text] [doi: 10.2196/11885] [Medline: 30664468]
18. 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 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
19. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: A meta-analysis. JAMA Pediatr 2017 May 01;171(5):461-469 [FREE Full text] [doi: 10.1001/jamapediatrics.2017.0042] [Medline: 28319239]
20. Chen CE, Harrington RA, Desai SA, Mahaffey KW, Turakhia MP. Characteristics of digital health studies registered in ClinicalTrials.gov. JAMA Intern Med 2019 Jun 01;179(6):838-840. [doi: 10.1001/jamainternmed.2018.7235] [Medline: 30801617]
21. Quinn CC, Swasey KK, Torain JM, Shardell MD, Terrin ML, Barr EA, et al. An mHealth diabetes intervention for glucose control: Health care utilization analysis. JMIR Mhealth Uhealth 2018 Oct 15;6(10):e10776 [FREE Full text] [doi: 10.2196/10776] [Medline: 30322839]
22. Wood CS, Thomas MR, Budd J, Mashamba-Thompson TP, Herbst K, Pillay D, et al. Taking connected mobile-health diagnostics of infectious diseases to the field. Nature 2019 Feb;566(7745):467-474 [FREE Full text] [doi: 10.1038/s41586-019-0956-2] [Medline: 30814711]
23. Hou C, Carter B, Hewitt J, Francisa T, Mayor S. Do mobile phone applications improve glycemic control (HbA1c) in the self-management of diabetes? A systematic review, meta-analysis, and grade of 14 randomized trials. Diabetes Care 2016 Nov;39(11):2089-2095. [doi: 10.2337/dc16-0346] [Medline: 27926892]
24. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-management education and support. J Diabetes Sci Technol 2017 Sep;11(5):1015-1027 [FREE Full text] [doi: 10.1177/1932296817713506] [Medline: 28560898]
25. Garabedian LF, Ross-Degnan D, Wharam JF. Mobile phone and smartphone technologies for diabetes care and self-management. Curr Diab Rep 2015 Dec;15(12):109 [FREE Full text] [doi: 10.1007/s11892-015-0680-8] [Medline: 26458380]
26. Auerbach AD. Evaluating digital health tools: Prospective, experimental, and real world. JAMA Intern Med 2019 Jun 01;179(6):840-841. [doi: 10.1001/jamainternmed.2018.7229] [Medline: 30801610]
27. Desveaux L, Shaw J, Saragosa M, Soobiah C, Marani H, Hensel J, et al. A mobile app to improve self-management of individuals with type 2 diabetes: Qualitative realist evaluation. J Med Internet Res 2018 Mar 16;20(3):e81 [FREE Full text] [doi: 10.2196/jmir.8712] [Medline: 29549070]
28. Hood M, Wilson R, Corsica J, Bradley L, Chirinos D, Vivo A. What do we know about mobile applications for diabetes self-management? A review of reviews. J Behav Med 2016 Dec;39(6):981-994. [doi: 10.1007/s10865-016-9765-3] [Medline: 27412774]
29. Chinese Diabetes Society. Guidelines for the prevention and control of type 2 diabetes in China (2017 Edition). Chin J Pract Intern Med 2018;38(4):292-344.

30. Austin PC. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. Commun Stat Simul Comput 2009 Apr 09;38(6):1228-1234. [doi: 10.1080/03610910902859574]
31. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70(1):41-55. [doi: 10.1093/biomet/70.1.41]
32. Ramos R, Comas-Cufí M, Martí-Lluch R, Balló E, Ponjoan A, Alves-Cabratosa L, et al. Statins for primary prevention of cardiovascular events and mortality in old and very old adults with and without type 2 diabetes: Retrospective cohort study. BMJ 2018 Sep 05;362:k3359 [FREE Full text] [doi: 10.1136/bmj.k3359] [Medline: 30185425]
33. Nelson DB, Rice DC, Niu J, Atay S, Vaporciyan AA, Antonoff M, et al. Long-term survival outcomes of cancer-directed surgery for malignant pleural mesothelioma: Propensity score matching analysis. J Clin Oncol 2017 Oct 10;35(29):3354-3362. [doi: 10.1200/JCO.2017.73.8401] [Medline: 28817374]
34. Xu Y, Wang L, He J, Bi Y, Li M, Wang T, 2010 China Noncommunicable Disease Surveillance Group. Prevalence and control of diabetes in Chinese adults. JAMA 2013 Sep 04;310(9):948-959. [doi: 10.1001/jama.2013.168118] [Medline: 24002281]
35. Mohan V, Mapari JA, Karnad PD, Mann JS, Maheshwari VK. Reduced diabetes mellitus-related comorbidities by regular self-monitoring of blood glucose: Economic and quality of life implications. Indian J Endocrinol Metab 2018;22(4):461-465 [FREE Full text] [doi: 10.4103/ijem.IJEM_216_17] [Medline: 30148089]
36. Altschul DM, Starr JM, Deary IJ. Cognitive function in early and later life is associated with blood glucose in older individuals: Analysis of the Lothian Birth Cohort of 1936. Diabetologia 2018 Sep;61(9):1946-1955 [FREE Full text] [doi: 10.1007/s00125-018-4645-8] [Medline: 29860628]
37. Nathan DM, Cleary PA, Backlund JC, Genuth SM, Lachin JM, Orchard TJ, Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Study Research Group. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. N Engl J Med 2005 Dec 22;353(25):2643-2653 [FREE Full text] [doi: 10.1056/NEJMoa052187] [Medline: 16371630]
38. Liese AD, Ma X, Reid L, Sutherland MW, Bell BA, Eberth JM, et al. Health care access and glycemic control in youth and young adults with type 1 and type 2 diabetes in South Carolina. Pediatr Diabetes 2019 May;20(3):321-329. [doi: 10.1111/pedi.12822] [Medline: 30666775]
39. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): Prospective observational study. BMJ 2000 Aug 12;321(7258):405-412 [FREE Full text] [doi: 10.1136/bmj.321.7258.405] [Medline: 10938048]
40. Laiteerapong N, Ham SA, Gao Y, Moffet HH, Liu JY, Huang ES, et al. The legacy effect in type 2 diabetes: Impact of early glycemic control on future complications (the Diabetes & Aging Study). Diabetes Care 2019 Mar;42(3):416-426. [doi: 10.2337/dc17-1144] [Medline: 30104301]
41. Khunti K, Ceriello A, Cos X, De Block C. Achievement of guideline targets for blood pressure, lipid, and glycaemic control in type 2 diabetes: A meta-analysis. Diabetes Res Clin Pract 2018 Mar;137:137-148. [doi: 10.1016/j.diabres.2017.12.004] [Medline: 29325774]
42. Jalving AC, Gant CM, Binnenmars SH, Soedamah-Muthu SS, Bakker SJ, Navis G, et al. Glycaemic control in the diabetes and Lifestyle Cohort Twente: A cross-sectional assessment of lifestyle and pharmacological management on HbA1c target achievement. Diabetes Obes Metab 2018 Oct;20(10):2494-2499 [FREE Full text] [doi: 10.1111/dom.13399] [Medline: 29862616]
43. Almutairi N, Hosseinzadeh H, Gopaldasani V. The effectiveness of patient activation intervention on type 2 diabetes mellitus glycemic control and self-management behaviors: A systematic review of RCTs. Prim Care Diabetes 2020 Feb;14(1):12-20. [doi: 10.1016/j.pcd.2019.08.009] [Medline: 31543458]
44. Vloemans AF, Eilander MM, Rotteveel J, Bakker-van Waarde WM, Houdijk EC, Nuboer R, et al. Youth with type 1 diabetes taking responsibility for self-management: The importance of executive functioning in achieving glycemic control: Results from the longitudinal DINO study. Diabetes Care 2019 Feb;42(2):225-231. [doi: 10.2337/dc18-1143] [Medline: 30552132]
45. Pulgarón ER, Sanders LM, Patiño-Fernandez AM, Wile D, Sanchez J, Rothman RL, et al. Glycemic control in young children with diabetes: The role of parental health literacy. Patient Educ Couns 2014 Jan;94(1):67-70 [FREE Full text] [doi: 10.1016/j.pec.2013.09.002] [Medline: 24091252]
46. Testa R, Bonfigli AR, Prattichizzo F, La Sala L, De Nigris V, Ceriello A. The "metabolic memory" theory and the early treatment of hyperglycemia in prevention of diabetic complications. Nutrients 2017 Apr 28;9(5):437 [FREE Full text] [doi: 10.3390/nu9050437] [Medline: 28452927]
47. Vervloet M, van Dijk L, de Bakker DH, Souverein PC, Santen-Reestman J, van Vlijmen B, et al. Short- and long-term effects of real-time medication monitoring with short message service (SMS) reminders for missed doses on the refill adherence of people with type 2 diabetes: Evidence from a randomized controlled trial. Diabet Med 2014 Jul;31(7):821-828. [doi: 10.1111/dme.12439] [Medline: 24646343]
48. Yin H, Liang S, He H, Liu XQ. Progress in therapeutic strategy for high glucose-induced metabolic memory. Prog Pharm Sci 2018;42(8):599-607.

49. Shen Y, Wang F, Zhang X, Zhu X, Sun Q, Fisher E, et al. Effectiveness of internet-based interventions on glycemic control in patients with type 2 diabetes: Meta-analysis of randomized controlled trials. J Med Internet Res 2018 May 07;20(5):e172 [FREE Full text] [doi: 10.2196/jmir.9133] [Medline: 29735475]

Abbreviations

FBG: fasting blood glucose
HbA_{1c}: glycated hemoglobin
LDL: low-density lipoprotein
mHealth: mobile health
P2BG: postprandial 2-hour blood glucose
VTD: variation trends of difference

Edited by G Eysenbach; submitted 07.07.19; peer-reviewed by S Ling, Y Hong, R Sun; comments to author 27.07.19; revised version received 11.09.19; accepted 09.02.20; published 11.03.20.

Please cite as:
Li J, Sun L, Wang Y, Guo L, Li D, Liu C, Sun N, Xu Z, Li S, Jiang Y, Wang Y, Zhang S, Chen L
A Mobile-Based Intervention for Glycemic Control in Patients With Type 2 Diabetes: Retrospective, Propensity Score-Matched Cohort Study
JMIR Mhealth Uhealth 2020;8(3):e15390
URL: http://mhealth.jmir.org/2020/3/e15390/
doi:10.2196/15390
PMID:32159518

Original Paper

Compliance and Utility of a Smartphone App for the Detection of Exacerbations in Patients With Chronic Obstructive Pulmonary Disease: Cohort Study

Juan Luis Rodriguez Hermosa[1,2], MD; Antonia Fuster Gomila[3], MD; Luis Puente Maestu[4], MD; Carlos Antonio Amado Diago[5,6], MD; Francisco Javier Callejas González[7], MD; Rosa Malo De Molina Ruiz[8], MD; Manuel E Fuentes Ferrer[9,10], MD; Jose Luis Álvarez Sala-Walther[1,2], MD; Myriam Calle Rubio[1,2], MD

[1]Pulmonology Department, Hospital Clínico San Carlos, Madrid, Spain
[2]Medical Department, School of Medicine, Universidad Complutense de Madrid, Madrid, Spain
[3]Hospital Universitario Son Llátzer, Palma de Mallorca, Spain
[4]Pulmonology Department, Hospital Universitario Gregorio Marañón, Madrid, Spain
[5]Pulmonology Department, Hospital Universitario Marqués de Valdecilla, Santander, Spain
[6]Medical Department, School of Medicine, Universidad de Cantabria, Santander, Spain
[7]Pulmonology Department, Complejo Hospitalario Universitario de Albacete, Albacete, Spain
[8]Pulmonology Department, Hospital Universitario Puerta de Hierro de Majadahonda, Madrid, Spain
[9]Departament of Preventive Medicine, Hospital Clínico San Carlos, Madrid, Spain
[10]Instituto de Investigación Sanitaria del Hospital Clínico San Carlos, Madrid, Spain

Corresponding Author:
Myriam Calle Rubio, MD
Pulmonology Department
Hospital Clínico San Carlos
Calle Profesor Martin Lagos
Madrid, 28040
Spain
Phone: 34 913303000 ext 3477
Email: mcallerubio@gmail.com

Abstract

Background: In recent years, mobile health (mHealth)–related apps have been developed to help manage chronic diseases. Apps may allow patients with a chronic disease characterized by exacerbations, such as chronic obstructive pulmonary disease (COPD), to track and even suspect disease exacerbations, thereby facilitating self-management and prompt intervention. Nevertheless, there is insufficient evidence regarding patient compliance in the daily use of mHealth apps for chronic disease monitoring.

Objective: This study aimed to provide further evidence in support of prospectively recording daily symptoms as a useful strategy to detect COPD exacerbations through the smartphone app, Prevexair. It also aimed to analyze daily compliance and the frequency and characteristics of acute exacerbations of COPD recorded using Prevexair.

Methods: This is a multicenter cohort study with prospective case recruitment including 116 patients with COPD who had a documented history of frequent exacerbations and were monitored over the course of 6 months. At recruitment, the Prevexair app was installed on their smartphones, and patients were instructed on how to use the app. The information recorded in the app included symptom changes, use of medication, and use of health care resources. The patients received messages on healthy lifestyle behaviors and a record of their cumulative symptoms in the app. There was no regular contact with the research team and no mentoring process. An exacerbation was considered reported if medical attention was sought and considered unreported if it was not reported to a health care professional.

Results: Overall, compliance with daily records in the app was 66.6% (120/180), with a duration compliance of 78.8%, which was similar across disease severity, age, and comorbidity variables. However, patients who were active smokers, with greater dyspnea and a diagnosis of depression and obesity had lower compliance (P<.05). During the study, the patients experienced a total of 262 exacerbations according to daily records in the app, 99 (37.8%) of which were reported exacerbations and 163 (62.2%)

were unreported exacerbations. None of the subject-related variables were found to be significantly associated with reporting. The duration of the event and number of symptoms present during the first day were strongly associated with reporting. Despite substantial variations in the COPD Assessment Test (CAT), there was improvement only among patients with no exacerbation and those with reported exacerbations. Nevertheless, CAT scores deteriorated among patients with unreported exacerbations.

Conclusions: The daily use of the Prevexair app is feasible and acceptable for patients with COPD who are motivated in their self-care because of frequent exacerbations of their disease. Monitoring through the Prevexair app showed great potential for the implementation of self-care plans and offered a better diagnosis of their chronic condition.

(JMIR Mhealth Uhealth 2020;8(3):e15699) doi:10.2196/15699

KEYWORDS

chronic obstructive pulmonary disease; mHealth; compliance; mobile phone

Introduction

Chronic obstructive pulmonary disease (COPD) places an enormous burden on health care systems. A substantial proportion of the cost is attributable to hospitalizations, mostly owing to acute exacerbations of respiratory symptoms [1]. Acute exacerbations of COPD (AECOPD) have important consequences for patients and health care providers; AECOPD cause a negative impact on health-related quality of life [2,3], a decline in pulmonary function [4], increased utilization of health care resources [5], and decreased survival [6,7].

Knowledge about COPD exacerbation frequency is important to assess the clinical risk [8-10]. The identification and correct assessment of COPD exacerbations is vital, given that it will strongly influence therapy success and the impact on patients' morbidity, mortality, and quality of life. Some individuals appear more susceptible to developing exacerbations and are termed frequent exacerbators or COPD exacerbator phenotypes [11,12]. Patients with frequent exacerbations are specifically targeted with more aggressive therapy and an action plan to help prevent exacerbations [8,9] and improve their quality of life [13,14]. In addition, the early identification of exacerbations by patients and early treatment may have effects on patient-reported outcomes [15]. Despite growing evidence supporting the importance of the identification and correct assessment of COPD exacerbations, less than one-third of exacerbations remain unreported to health care professionals. Unreported exacerbations are common and important events. Several studies through questionnaires showed that nearly half of all exacerbations remain unreported [16,17]. Exacerbations that are unreported and untreated by health care professionals are associated with worsening in the quality of life [3,18,19] and an increased risk of subsequent hospitalization [20,21]. Failure to seek medical attention has consequences. There is a need for new strategies to capture symptom-based exacerbations and thus provide better management of COPD.

In recent years, health-related apps running on mobile devices such as smartphones and tablets, known as mobile health (mHealth) apps, have been developed to help manage chronic diseases [22]. Current apps aid patients in managing their chronic disease, aid patients in adopting a healthy lifestyle (good nutrition, exercise, and smoking cessation), and can aid in providing a better quality of life for patients. In patients with chronic disease characterized by exacerbations, such as COPD, apps may allow tracking and even alert patients and health care professionals about suspected disease exacerbations, thereby facilitating self-management and prompt intervention [23-25]. However, maintaining continuous use is still a challenge. The use of apps to capture the reality of subjects' lives in chronic disease is quite problematic [26,27]. The complex issues around compliance, and the relationship between symptoms and behavior, has led to the need for more research examining the parameters that contribute to daily compliance [28,29]. There is insufficient evidence regarding patient compliance in the daily use of mHealth apps for chronic disease monitoring and the determinants that contribute to this compliance. We were especially interested in these issues in the context of COPD, where little research has been done on recording symptoms daily using mHealth apps.

This study aimed to provide further evidence in support of the hypothesis that prospectively recording daily symptoms is a useful tool to monitor and help correctly assess COPD exacerbations and the clinical risk of COPD based on the patient's daily self-reporting of symptoms using Prevexair, a simple smartphone app, in which the patient records their daily symptoms and which offers general recommendations. As there is not enough information based on previously reported results, this study was proposed as a pilot study to generate information.

This paper has provided data about compliance in the daily use of a mHealth app for the long-term monitoring of patients with COPD without a mentoring process or regular phone calls. In addition, this paper has analyzed the frequency and characteristics of AECOPD recorded via the smartphone app, Prevexair.

Methods

Study Patients

Patients were recruited in outpatient respiratory clinics from 6 tertiary referral hospitals in Spain between November 2016 and March 2018. The inclusion criteria were as follows: aged above 40 years, having a history of smoking (≥10 pack-years), a diagnosis of COPD confirmed by postbronchodilator spirometry with a forced expiratory volume in one second (FEV_1) to forced vital capacity ratio of less than 0.7 in the stable phase of the disease, having a history of at least two exacerbations treated with oral corticosteroids or antibiotics or having been hospitalized at least once for exacerbation in the past 12 months, owning a smartphone, and having the cognitive and motor ability to operate a smartphone. Patients were excluded if they had

other significant respiratory diseases or if they reported an exacerbation during the run-in period. Ethical approval was obtained from the Ethics Committee at the Hospital Clínico San Carlos (Madrid, Spain; internal code 14/124-E), and all patients gave their written informed consent before inclusion.

Study Design and Patient Evaluation

This was a multicenter, prospective cohort study with a 2-week run-in period followed by a 6-month follow-up period. The study visits were scheduled as follows: before the 2-week run-in period (selection visit), after the run-in period (inclusion visit), during the follow-up period at 3 months (visit 1), and at 6 months (visit 2). The run-in period was used to make sure patients were stable. Patient assessment included a complete medical history (height, weight, smoking history, drug history, diagnosis of depression and/or anxiety, and other comorbid conditions), spirometry, health-related quality of life using the COPD Assessment Test (CAT), dyspnea using the modified Medical Research Council (mMRC) questionnaire, and the number of moderate/severe exacerbations in the last year.

Data Collection and Monitoring

The app was developed for iOS and Android systems by using Virtual Ware. The app is available for installation on a mobile device. At recruitment, during selection visit, the Prevexair app was installed on patients' smartphones, and they were instructed how to use the app and received instructions to record their daily respiratory symptoms in the app, once under supervision. Thereafter, in the inclusion visit, their ability to use the app was reviewed. No problems were recorded regarding the use of the mobile app.

The information recorded in the app included symptom changes, use of medication, and use of health care resources. Figures 1 and 2 show several screenshots of the Prevexair app in a smartphone. The following symptoms were included in the app: dyspnea, sputum color and amount, wheeze, cough, colds, and sore throat. The symptom questions had dichotomous response options, with a positive response indicating that the symptom was worse than at baseline. In addition, patients were instructed to use the app to record whether they increased their inhalation medication and started corticosteroids or antibiotics as well as any medical assistance. Daily data entries were made in the evening by setting a reminder alarm. If they forgot, patients were only allowed to enter the data for the 3 previous days. Patients automatically sent these records to a central server to be monitored in real-time by the research team. App users received messages on healthy lifestyle behaviors and disease education, information about their medication, task notifications, and a record of their cumulative symptoms in a graph through the app. Occasional contact was made, but only to solve minor technical problems. No regular contact was established, and no mentoring process was implemented to increase compliance. Participants were informed that the data sent would only be consulted by the research group; however, their physicians were blinded; they were not informed about the Prevexair app records. Thus, if they felt ill, they should contact their regular physicians for advice as usual. Decisions to change treatment or to go to the hospital will be made according to usual practice upon orders from the primary care or the respiratory specialist treating the patient, but these decisions were not based on the information provided during the study.

Additional information gathered at each scheduled visit during a clinically stable period included changes in medication, current smoking habits, CAT, and use of health care resources. During the visit at 6 months, the level of satisfaction with the app was evaluated on a scale from 0 to 10. The medical staff at the visit were blinded and not informed about the Prevexair app records.

Figure 1. Screenshots of the Prevexair app in a smartphone: initial screens.

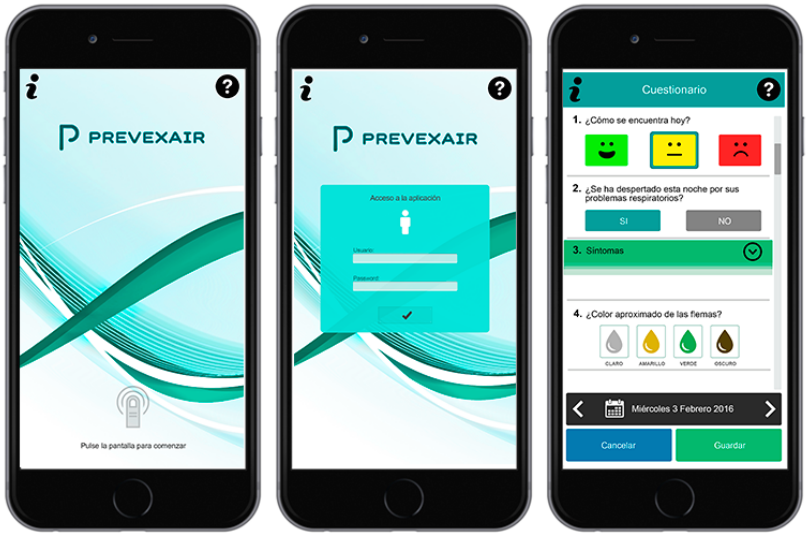
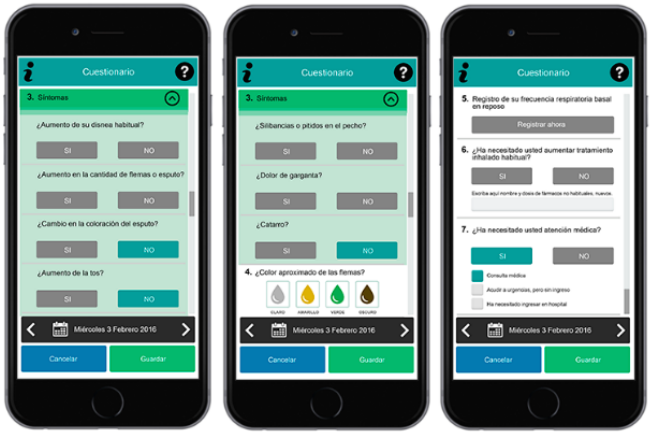


Figure 2. Screenshots of the Prevexair app in a smartphone: questionnaire.



Data Management

To analyze data recording compliance in the app, there were 180 *available days* for each participant, including the periods of hospitalization. Two measurements of compliance were defined: overall compliance (percentage of available days that data were recorded in the app) and duration of compliance (days elapsed from the first to last app entry, as a percentage of available days).

An exacerbation was defined as an increase in respiratory symptoms for 2 consecutive days, with at least one major symptom (dyspnea, sputum purulence, or sputum volume) and either another major or minor symptom (wheeze, cold, sore throat, or cough). The first day a symptom worsened was defined as the day of onset of the exacerbation, according to the previously validated criteria [30]. Symptom counts were obtained by adding each increased respiratory symptom recorded in the Prevexair app on the first day of the exacerbation.

Exacerbation duration was defined as the number of days after onset that worsened symptoms persisted. The last day of recorded worsened symptoms before 2 consecutive symptom-free days was defined as the end of the exacerbation. Patients had to be symptom-free for ≥7 days before a new exacerbation onset was defined. Exacerbation recovery could not be determined if patients failed to record symptoms or continuously recorded symptoms for more than 30 days after onset.

An exacerbation was considered reported if medical attention was sought through scheduled or unscheduled doctor visits, emergency department visits, or hospital admissions. It was considered unreported if it was not medically reported. An exacerbation was considered treated if there was a change in at least one medication (ie, antibiotics, corticosteroids, or bronchodilators) for the worsened symptom.

To analyze the impact of exacerbation on the health status, the patients were categorized into four exclusive groups according to the reporting status for exacerbations: no exacerbation, unreported exacerbation only, reported exacerbation only, and mixed unreported (at least one) and reported (at least one) exacerbations. For this analysis, those patients with less than 60% overall compliance were excluded.

Statistical Analysis

Qualitative variables were summarized by their frequency distribution and quantitative variables by their mean and standard deviation. Continuous nonnormally distributed variables were summarized by median and interquartile range (IQR P25-P75).

The association between the quantitative variable level of compliance and patient characteristics was evaluated with the nonparametric Mann-Whitney test for two independent groups or the Kruskal-Wallis test for more than two groups.

The association between each independent variable (baseline patient characteristics and event characteristics) and the dependent variable type of exacerbation (unreported or reported) was assessed by calculating the crude odds ratio via a multilevel logistic regression analysis. The multilevel analysis included two levels: the event level (level 1) and the patient level (level 2). A multivariable, multilevel logistic model was fitted to evaluate the independent effect of the selected variables. Candidate predictors with $P<.10$ in the univariate analysis were accepted for inclusion in the multivariate analysis.

In the study of the relationship between exacerbations and quality of life, quantitative variables were compared between the four groups in the study using the Kruskal-Wallis test and qualitative variables were compared using the chi-square test.

All analyses were performed using STATA 15.0 software (StataCorp LLC). Statistical significance was assumed at $P<.05$.

Results

Demographic and Clinical Information

Of the 126 patients recruited, 10 patients were excluded because they had one exacerbation at enrollment. A total of 116 patients were eligible for analysis; 21.6% (25/126) of participants were women and 13.8% (16/116) were active smokers. The baseline characteristics of the analyzed cohort are reported in Table 1. The mean (SD) FEV_1 was 1.19 (0.47) L, and the percent predicted FEV_1 was 44.6 (16.2). In total, 53.4% (62/116) of participants had experienced at least one severe exacerbation in the last year, and 76.7% (84/116) of participants had a degree of dyspnea ≥2 mMRC.

Table 1. Baseline characteristics of the study population.

Characteristics	Values
Patients, n	116
Gender (male), n (%)	91 (78.4)
Age (years), mean (SD)	66.51 (8.14)
Active smokers, n (%)	16 (13.8)
Smoking pack-years, mean (SD)	44.1 (23.60)
BMI (kg/m^2)	
Mean (SD)	27.30 (4.96)
≤21, n (%)	9 (7.8)
Number of comorbidities	
Mean (SD)	2.62 (1.41)
≥3, n (%)	58 (50.0)
Dyspnea (mMRC[a]), n (%)	
0-1	27 (23.2)
≥2	89 (76.7)
CAT[b] questionnaire, mean (SD)	14.10 (6.13)
Chronic bronchitis, n (%)	61 (52.6)
Chronic colonization, n (%)	2 (1.7)
History of asthma, n (%)	7 (6.0)
Post-BD FEV_1 (%)[c], mean (SD)	44.62 (16.23)
Post-BD FEV_1 (mL), mean (SD)	1192.11 (477.16)
Number of severe-moderate exacerbations[d] in the last year, median (IQR P25-75)	3(2-3)
Number of moderate exacerbations in the last year ≥2, n (%)	65 (56.0)
Number of severe exacerbations in the last year, median (IQR P25-75)	1 (0-2)
Number of severe exacerbations in the last year ≥1, n (%)	62 (53.4)
Drug treatment for COPD[e], n (%)	
LAMA[f] monotherapy	6 (5.2)
LAMA-LABA[g] combination	37 (31.9)
LABA+ICS[h] combination	6 (5.2)
Triple therapy[i]	67 (57.7)
Long-term oxygen therapy, n (%)	36 (31.0)
Chronic antibiotics, n (%)	1 (0.9)

[a]mMRC: modified Medical Research Council.

[b]CAT: Chronic Obstructive Pulmonary Disease Assessment Test.

[c]Post-BD FEV_1 %: postbronchodilator FEV_1 percent predicted.

[d]Severe exacerbations refers to exacerbations requiring hospitalization; moderate exacerbations refer to exacerbations requiring outpatient management with antibiotics and/or corticosteroids systemic.

[e]COPD: chronic obstructive pulmonary disease.

[f]LAMA: long-acting antimuscarinic agent.

[g]LABA: long-acting beta-2 agonist.

[h]ICS: inhaled corticosteroid.

[i]Triple therapy: LABA+LAMA+ICS.

Daily Compliance

The 116 patients recorded data in the Prevexair app for a median of 178 (IQR 130-180) days, while the median number of records per patient was 120 (IQR 61-164).

Overall compliance in recording data daily in the app was 66.6% (120/180), with a compliance duration of 78.8%. Compliance is reported in Table 2. Overall and compliance duration rates were similar across the disease severity, age, and comorbidity variables. However, patients who are active smokers, with a higher mMRC functional dyspnea score and a diagnosis of depression and obesity, had lower overall compliance and duration of compliance ($P<.05$). Female participants also had a lower duration of compliance. Furthermore, participants with a higher degree of satisfaction with the app had better overall compliance and duration of compliance ($P<.05$).

The percentage of patients who discontinued their use of the app was 6% during the first month and 8.6% during the second and third months, and 12.9% of patients abandoned the app during the last 3 months. The median (P25-75) level of satisfaction with the app was 10 (8-10).

Table 2. Daily compliance by clinical variable.

Characteristics	Participants, n	Overall compliance[a], median (IQR)	*P* value	Compliance duration[c], median (IQR)	*P* value
All subjects	116	66.6 (33.8-91.1)	N/A[b]	78.8 (51.5-94.9)	N/A
Sex			.10		.03
Male	91	71.6 (35.5-93.3)		82.7 (56.7-95.6)	
Female	25	53.3 (31.9-79.1)		74.4 (42.1-85.2)	
Age (years)			.79		.88
<65	44	63.8 (34.8-90.5)		78.1 (45.1-95.5)	
≥65	72	67.5 (33.8-91.9)		79.4 (56.7-92.7)	
BMI (kg/m^2)			.02		.01
25-26.9	40	78.6 (42.2-95.4)		80 (59.8-95.4)	
27-29.9	50	72.2 (42.9-92.2)		85.7 (64.7-96.3)	
≥30	26	40.5 (11.2-75.5)		54.5 (32.0-85.6)	
Smoking status			.03		.02
Active smoker	16	40.5 (7.9 -72.6)		54 (30.4-84.6)	
Former smoker	100	71.6 (40.0-91.9)		81.6 (56.7-95.0)	
Number of comorbidities			.88		.37
<3	58	64.1 (40.6-89.7)		75.3 (53.8-93.1)	
≥3	58	72.2 (31.1-92.3)		84.2 (45.3-95.8)	
Chronic bronchitis			.57		.21
Presence	61	64.4 (32.2-89.4)		77.9 (44.9-92.0)	
Absence	55	68.3 (39.4-95.0)		85.5 (60.5-96.2)	
History of asthma			.86		.66
Presence	12	59.1 (35.9-89.8)		70.2 (55.0-90.1)	
Absence	104	67.5 (33.8-91.1)		80 (50.9-94.9)	
Chronic colonization			.57		.21
Presence	61	64.4 (32.2-89.4)		77.9 (44.9-92.0)	
Absence	55	68.3 (39.4-95.0)		85.5 (60.5-96.2)	
Dyspnea (mMRC[d])			.08		.04
0-1	27	80.5 (45-98.8)		87.9 (61.3-100)	
≥2	89	63.8 (32.2-90.5)		75.6 (47.1-91.9)	
CAT[e]questionnaire			.76		.97
<10	31	61.6 (41.6-92.7)		76.5 (50-96.7)	
≥10	84	70 (33.3-90.5)		80.5 (54.5-92.6)	
Anxiety			.13		.34
Presence	21	88.3(47.5-93.6)		88.3 (52.3-94.3)	
Absence	95	63.8 (33.3-90.5)		75.7 (51.3-94.9)	
Depression			.008		.01
Presence	7	12.2 (3.3-53.3)		32.1 (9.2 -87.2)	
Absence	109	71.6 (39.7-91.6)		80.5 (56.6-95.0)	
Post-BD FEV_1 (%)[f]			.78		.65
≥50	37	71.6 (36.6-88.8)		81.6 (46.5-91.8)	
<50	79	64.4 (33.3-92.2)		77.9 (53.8-95.1)	

Characteristics	Participants, n	Overall compliance[a], median (IQR)	*P* value	Compliance duration[c], median (IQR)	*P* value
Number of moderate/severe exacerbations[g] in the last year			.30		.74
<2	11	51.6 (7.2- 86.6)		80.6 (61.1-94.2)	
≥2	105	68.3 (37.5-91.1)		78.8 (50.6-94.9)	
Drug treatment for COPD[h]			.42		.64
No triple therapy	49	71.6 (47.2-90.0)		78.8 (55.0-90.9)	
Triple therapy[i]	67	66.6 (25.5-92.2)		80.2 (45.0-95.1)	
Long-term oxygen therapy			.21		.10
Treated	36	80.5 (33.1-95.0)		87.1 (71.5-95.0)	
Not treated	80	63.3 (33.8-90.2)		73.6 (45.8-92.3)	
Satisfaction score			.10		.20
<10	50	63.8 (32.5-90.5)		73.6 (47.5-90.8)	
≥10	54	78.6 (44.0-93.7)		85.6 (55.4-96.2)	

[a]Compliance is expressed as median percentage (number of days completed/total number of days available for completion). Overall compliance: percentage of days in the entire study period (180 days) in which the app was used daily.

[b]Not applicable.

[c]Compliance duration: days elapsed from first daily entry to last, as the percentage of days available.

[d]mMRC: modified Medical Research Council.

[e]CAT: Chronic Obstructive Pulmonary Disease Assessment Test.

[f]Post-BD FEV_1 %: postbronchodilator FEV_1 percent predicted.

[g]Severe exacerbations refer to exacerbations requiring hospitalization; moderate exacerbations refer to exacerbations requiring outpatient management with antibiotics and/or corticosteroids systemic.

[h]COPD: chronic obstructive pulmonary disease.

[i]Triple therapy: long-acting beta-2 agonists + long-acting antimuscarinic agents + inhaled corticosteroids.

Exacerbations

During the study, patients experienced a total of 262 cases of symptom worsening, meeting the definition of exacerbation according to daily records in the app. The overall estimated rate of exacerbations recorded in the app was 2.25 (1.66) per person every 6 months. Of 116 patients, 18 (15.5%) had no events, 26 (22.4%) had one event, 25 (21.6%) patients had 2 events, and 47 (41.6%) patients had more than 2 events during the 6 months. Of 262 cases, 99 (37.8%) were reported exacerbations and 163 (62.2%) were unreported exacerbations.

Table 3 presents the characteristics of reported and unreported exacerbations and their relationship with the probability of reporting. In general, reported exacerbations were longer and had more symptoms: among those exacerbations with 2 symptoms present at onset, only 15.2% were reported, whereas 46.5% of those with 4 or more symptoms were reported. In reported exacerbations, sputum color (54.5% vs 22.7%; *P*<.001) and cough (74% vs 60.1%; *P*<.02) were also more common. Reporting was related to the duration and number of worsened symptoms as well as the type of symptoms present when symptoms worsened.

Of the 163 unreported exacerbations, 76 (46.6%) were treated, but all were self-managed by the patient with an increase in bronchodilators in 57 events (35.0%), only antibiotics in 10 events (6.1%), and only oral corticosteroids in 9 (5.5%) events. Of the 99 reported exacerbations, all were treated: 15 (15%) events only with an increase in bronchodilators and the majority with only antibiotics (47/99, 47.5%) or with both oral corticosteroids and antibiotics (28/99, 28.3%). With regard to recorded health care utilization for exacerbation, 79.7% of exacerbations led to unscheduled contact, 2.2% led to emergency department visits, and 18.1% of exacerbations resulted in hospitalization.

Table 3. Characteristics of unreported and reported exacerbations and the relationship between event characteristics and the likelihood of reporting an exacerbation.

Characteristics of exacerbations	Global	Unreported	Reported	Odds ratio (95% CI)	*P* value
Exacerbations, n (%)	262 (100.0)	163 (62.2)	99 (37.8)	N/A[a]	N/A
Duration of worsened symptoms, day median (P25-75)	6 (4-9)	5 (3-8)	8 (6-11.2)	1.17 (1.08-1.27)	<.001
Total number of key symptoms, mean (SD)	3.12 (1.09)	2.91 (1.04)	3.47 (1.09)	1.88 (1.35-2.63)	<.001
Type of symptoms, n (%)					
Dyspnea	137 (52.3)	87 (53.4)	50 (50.5)	0.95 (0.49-1.00)	.9
Sputum amount	177 (67.6)	106 (65.0)	71 (71.7)	1.48 (0.72-3.05)	.28
Sputum color	91 (34.7)	37 (22.7)	54 (54.5)	8.39 (3.31-21.22)	<.001
Cough	172 (65.6)	98 (60.1)	74 (74.7)	2.30 (1.13-4.66)	.02
Wheeze	75 (28.6)	45 (27.6)	30 (30.3)	1.31 (0.62-2.75)	.47
Sore throat	57 (21.8)	37 (22.7)	20 (20.2)	0.69 (0.29-1.63)	.49
Cold	110 (41.9)	65 (39.9)	45 (45.5)	1.30 (0.64-2.61)	.46
Severity by number of symptoms, n (%)				2.93 (1.25-6.85)	<.001
2	85 (32.5)	68 (41.7)	17 (17.1)		
3	96 (36.6)	60 (36.8)	36 (36.4)		
4 or more	81 (30.9)	35 (21.5)	46 (46.5)		

[a]Not applicable.

Baseline Characteristics of Patients by Exacerbation Category

Table 4 shows that patients with reported exacerbations present similar baseline characteristics compared with patients with unreported exacerbations, although patients with reported exacerbations had more dyspnea, severe disease FEV_1%, and anxiety. Reporting did not appear to be related to patient characteristics.

Table 4. Baseline characteristics of patients by exacerbation category.

Characteristics	Unreported exacerbations	Reported exacerbations	Odds ratio (95% CI)	*P* value
Gender (male), n (%)	123 (75.4)	82 (83)	1.61 (0.61-4.29)	.33
Age (years)				
Mean (SD)	65.5 (8.7)	67.6 (8.3)	1.04 (0.58-4.68)	.08
≥65, n (%)	102 (62.5)	68 (69)	1.60 (0.67-3.81)	.28
Active smokers, n (%)	27 (16.6)	12 (12)	1.68 (0.53-5.26)	.37
BMI (kg/m^2), mean (SD)	26.6 (5.0)	26.8 (4.6)	1.02 (0.94-1.11)	.55
Number of comorbidities; mean (SD) ≥3, n (%)	68 (41.7)	55 (56)	2.11 (0.94-4.73)	.07
Depression, n (%)	8 (4.9)	3 (3)	0.55 (0.07-3.97)	.56
Anxiety, n (%)	25 (15.3)	26 (26)	2.77 (0.99-7.70)	.05
Dyspnea (mMRC[a]) ≥2, n (%)	129 (79.1)	87 (88)	2.73 (0.88-8.41)	.08
CAT[b] questionnaire score ≥10, n (%)	115 (70.5)	80 (81)	2.28 (0.88-5.94)	.09
Post-BD FEV_1 %[c] predicted <50%, n (%)	114 (69.9)	74 (75)	1.58 (0.64-3.91)	.32
Number of moderate-severe exacerbations in the last year ≥2, n (%)	149 (91.4)	96 (97)	4.06 (0.69-23.72)	.12

[a]mMRC: modified Medical Research Council.

[b]CAT: Chronic Obstructive Pulmonary Disease Assessment Test.

[c]Post-BD FEV_1 %: postbronchodilator FEV_1 percent predicted.

Predictors of Reporting an Exacerbation

Table 5 shows the relationship between exacerbation characteristics and the subject and the likelihood of reporting an exacerbation. None of the subject-related variables were found to be significantly associated with reporting. The duration of the event and number of symptoms present during the first day were strongly associated with reporting.

Table 5. The relationship between exacerbation characteristics and the subject and the likelihood of reporting an exacerbation.

Characteristics	Odds ratio (95% CI)	*P* value
Duration of worsened symptoms	1.15 (1.06-1.25)	<.001
Mean number of key symptoms	1.75 (1.20-2.56)	.003
Number of comorbidities		.45
≥3	1.50 (0.51-4.41)	
<3	1	
CAT[a] questionnaire score		.52
≥10	1.49 (0.44-5.00)	
<10	1	
Dyspnea (mMRCM[b])		.42
≥2	1.75 (0.44-6.99)	
<2	1	
Anxiety		.20
Present	2.43 (0.62-9.50)	
Not present	1	

[a]CAT: Chronic Obstructive Pulmonary Disease Assessment Test.

[b]mMRC: modified Medical Research Council.

Impact of Exacerbations on Health Status

On average, CAT scores were worse at the end of the study period (6 months). The median (P25-75) change in CAT score was 1 (−3-4). Table 6 shows the distribution of CAT score changes between the inclusion and 6-month visits stratified by the presence and type of exacerbation recorded in the app among 69 patients who had more than 60% overall compliance. Of the 69 patients analyzed, only 6 (9%) did not have exacerbations, 19 (27%) had only unreported exacerbations, 18 (26%) had only reported exacerbations, and 26 (38%) had mixed exacerbations. Despite substantial variation in the CAT score, there was improvement among patients with no exacerbations and those with only reported exacerbations. CAT scores deteriorated in patients with unreported exacerbations. This deterioration was highest in patients who had at least one unreported exacerbation and at least one reported exacerbation. There was a difference between those with only unreported exacerbations and those with only reported exacerbations ($P<.05$). Patients in the mixed exacerbation group had statistically worse deterioration ($P<.01$) of the CAT score than those with only reported exacerbations. Deterioration of the CAT score was clinically significant (an increase of 2 or more) in 44% (8/18) of patients who did not report any of their exacerbations compared with 23% (4/18) of those with only reported exacerbations and 73% (19/26) of those with mixed exacerbations.

Table 6. Change in health status between inclusion and 6-month visits according to the presence and type of exacerbation during the study as recorded in the app.

Change in health status	Stable disease[a]	Unreported[b]	Reported[c]	Mixed[d]	*P* value
Subjects, n (%)	6 (9)	19 (28)	18 (26)	26 (38)	N/A[e]
Change in CAT[f] score, median (P25-75)	−3 (−3.5-3)	1 (−2.2-6.2)	−2 (−7-1.5)	3 (0-5.2)	<.001
Patients with change in CAT score ≥2, n (%)	1 (20)	8 (44)	4 (23)	19 (73)	<.001

[a]No exacerbation between inclusion and 6-month visits.

[b]Only unreported exacerbation(s) between inclusion and 6-month visits.

[c]Only reported exacerbation(s) between inclusion and 6-month visits.

[d]At least one unreported exacerbation and one reported exacerbation between inclusion and 6-month visits.

[e]Not applicable.

[f]CAT: Chronic Obstructive Pulmonary Disease Assessment Test.

Discussion

Principal Findings

This study provides information about the long-term, consistent use of an mHealth app, Prevexair, to record daily symptoms and detect exacerbations in high-risk patients with COPD, as well as to determine the characteristics of the detected exacerbations and the determinants of reporting them.

The mHealth app market is booming and will continue to grow substantially over the next few years. The growing availability of health apps and the increasing number of patients using smartphones and tablets will encourage health care professionals to incorporate apps into their management plans for patients with chronic disease. This is a step toward ubiquitous health care, thereby allowing patients with chronic disease to self-manage their condition by providing them support to monitor and interpret their own data using mobile devices.

COPD is a highly prevalent disease, occurring in 10% of the population between the ages of 40 and 80 years [31]. It is a progressive disease that is frequently associated with a high rate of morbidity and mortality and is currently the fifth leading cause of death in Spain [32,33]. It is currently included in the priority plans for health care systems [34] owing to its association with a significant demand for care because of its high complexity and frequent decompensations [35]. COPD is one of the main reasons for medical consultations and the use of health care resources, both in primary and specialized care. In Spain, the disease accounts for 10% to 12% of all primary care visits, 35% to 40% of pulmonology consults, and 7% of hospitalizations [36,37]. These characteristics of COPD force us to make a change in the care model, focusing on monitoring the disease and giving the patient a part of the responsibility in managing their disease through the use of information and communication technology as a tool that has been proven useful in the self-care and monitoring of patients with COPD to detect decompensations of the disease.

Previous Studies

Research has shown that effective management of COPD through integrated care systems, mHealth apps, and other technology has the potential to both benefit the patient and reduce exacerbation costs in the long-term management of the disease [38,39]. Several studies on action plans focusing on the early identification of exacerbations by patients and the implementation of an action plan have shown effects on health care utilization as well as on patient-reported outcomes [40-42]. The telemonitoring of a patient's condition, symptoms, and behavior (adherence to medication and physical activity) through mHealth apps may be useful in identifying and correctly assessing COPD exacerbations, reducing the number of unreported exacerbations, and allowing the implementation of self-management. Providing the patient with the right care at the right time is crucial and can have a decisive impact on handling the long-term condition to prevent exacerbations and improve the quality of life in patients with COPD. As a result, mHealth apps are extensively used in health services and patient education. Indeed, the UK Department of Health has recommended that apps be *prescribed* as part of the care for long-term conditions [43]. However, there are few published studies addressing daily compliance in mHealth apps and what factors influence compliance.

The results of our study show a high rate of daily use of the app, Prevexair, although there was no contact between the research team and the patient after initiation and no strategy was implemented to continue using the app.

In COPD, little research has been done on diary-keeping, even though diaries have been widely used in studies and clinical trials. In an open, observational study, only 41% of participants achieved 80% compliance using paper diaries that were collected weekly and entered electronically [44]. The compliance was higher (53% in 12 months) in another study owing to a mentoring process with regular phone calls [45].

Interpretation of Novel Findings

In our study, the level of satisfaction with the functionality of the Prevexair app was high. The patients quickly learned how to use the app during the inclusion visit and regularly entered data to record their symptoms and medicine use over 6 months, although they did not make decisions based on information provided by the app during the study. Decisions to change treatment or go to the hospital were made according to usual practice upon orders from the primary care or respiratory specialist treating the patient. Studies that have evaluated feedback from users regarding the functionality and usability of a mobile phone app show us that simplicity and motivation, not age, seem to be the key factors for accepting and using health apps [46]. With regard to these determinants of use, it is worth mentioning that the patients who participated in our study were motivated in their self-care because of frequent decompensations of their disease, with hospitalization for COPD occurring in more than half of the patients evaluated. There is evidence that participants will tolerate the burden of diary-keeping if they feel it will help them [47,48]. In this regard, it should be mentioned that the Prevexair app sent messages about healthy lifestyle behaviors and information about the patient's medication, task notifications, and a record of their cumulative symptoms in a graph through the app, which has been found to have a compliance advantage. Studies show that users value being in charge of their health and keeping track of their progress [49]. However, self-motivation to record data over a longer period can be a challenge without the involvement of a health care professional [50,51]. There is also evidence that participants in research studies will take on additional burdens for altruistic reasons unrelated to their chronic illness [52].

Factors Associated With Compliance

With regard to the factors related to continued daily use of the app, in our study, we did not find any differences according to age or sex, comorbidity burden, or disease severity. These results are consistent with other studies that showed that compliance rates were similar across the demographic variables of sex and disease severity [45,46]. Several studies have shown that there is no general correlation between diary-keeping and symptom severity, demographic or clinical characteristics, treatment, or activity [45,53].

We have provided new data on use and adherence to a mobile phone app for COPD. The compliance is not affected by demographic factors or disease severity, while clinical or physiological characteristics, such as actively smoked, higher BMI, or were diagnosed with depression, do seem to influence diary use. Nevertheless, a limitation to bear in mind is that other factors related to adherence such as health literacy, prior use of apps, and level of school education could not be evaluated as they were not available.

Simplicity and motivation seem to be the key factors for accepting and using mobile phone apps. However, each user has different needs, so it is important to be able to personalize the app to the patient's preferences. So, in the patients where we identify factors linked to lower adherence, it is important to offer specific messages such as exercise tracking, monitoring of weight, food intake, and help for tobacco cessation. In addition, personalized self-management plans could be updated according to patients' needs. Other functionality of interest can be email messaging or any type of communication with health care providers.

Detection of Exacerbations

Regarding the detection of exacerbations by recording symptoms in the app, it is necessary to highlight the high rate of daily exacerbations and that 62% of the events recorded in the app were not reported and most were not treated. This result is similar to results of earlier studies in London [16,19] and Canada [17]. These results support the management of COPD through mHealth apps and other technology, considering that COPD is a highly symptomatic disease, but patients may not recognize small day-to-day variations in their pulmonary symptoms. Lack of symptom awareness and the rate of symptom worsening make the daily monitoring of patients with COPD an attractive and beneficial approach to detect patients with frequent exacerbations and to carry out more aggressive therapy and implement preventive measures. In addition, these technologies would allow health care professionals to monitor patients and offer opportunities for an intervention to improve outcomes. They could view patients' data consistently, and not only periodically at the outpatient clinics. Furthermore, they could provide the patient with information to implement self-care and early treatment plans for COPD decompensations.

In our study, although unreported exacerbations tend to be milder (with a lower number of symptoms and shorter duration of exacerbation), these unreported exacerbations have a clinically relevant negative impact on quality of life (CAT questionnaire) and result in a change in self-administered treatment by the patient in a large number of cases. Patients who did not report their exacerbation were more likely to experience worsening of their health status compared with those who reported exacerbations or those with a stable disease. This may suggest that unreported exacerbations may thus represent an unmet health care need. These results are consistent with other studies, which have shown that unreported exacerbations, despite being associated with less symptom worsening than reported exacerbations, have an important medium to long-term impact on patients' quality of life [15,18,19]. Failure to seek medical attention may have consequences for both the patient and health care system.

In our study, the characteristics of the exacerbations were the strongest predictors of reporting. Although there was no direct measure of exacerbation severity, the total number of symptoms at onset and duration of the exacerbation were predictors of reporting. The symptoms associated with reporting an exacerbation in this study were cough and change in sputum color. Sputum color was identified as one of the key determinants of health care utilization in a study looking at patient perspectives on exacerbations [54]. Although in our study no subject-related variables were found to be significantly associated with reporting, patients who reported their exacerbation had more dyspnea, anxiety, and more spirometric obstruction compared with those who did not report exacerbations. Other studies have also found that patients with a lower FEV_1 were more likely to seek medical attention [17,55]. This is also consistent with the observation showing that physician-rated exacerbation severity correlates with the severity of the underlying disease [56]. The finding that patients with anxiety were more likely to seek medical attention may be explained by the fact that it plays a role in symptom development and patient behavior and should be considered a potentially important predictor. In other studies, psychological factors have been found to be related to the reporting of respiratory symptoms [54,57].

The identification and correct assessment of COPD exacerbations is important to assess clinical risk and disease control, a goal that is key especially in patients who appear more susceptible to developing exacerbations and are termed frequent exacerbators, similar to our study population, in which monitoring through the app, Prevexair, can be more beneficial.

The app was developed for better lifestyle management for patients with COPD and also to improve monitoring and follow-up by their physicians. In the future, we would like to analyze the usefulness of the app, Prevexair, for physicians during clinician visits for identification of COPD exacerbations and for the correct assessment of clinical risk of COPD, as the app offers the possibility to regularly record relevant health data of a patient's condition and symptoms. A strategy that could prove useful as several studies suggest that close to half of all exacerbations remain unreported. The unreported exacerbations and consequent lack of treatment by a health care professional were associated with worsening quality of life and increased risk of hospitalization.

Potential Strengths and Limitations

A limitation that must be considered in the interpretation of the results is that it provides us information about how app users will perform within the context of a research study as the benefit perceived by the patient is a determining factor in the motivation to use the app. During the study, no decisions were made based on the information recorded in the app. Participants were informed that if they felt ill, they should contact their regular physicians for advice. Other limitations of this study are that we have not evaluated other factors that seem to influence daily compliance and affect both the health status and access to health care, such as socioeconomic status, impact on activities of daily

living, and education level. However, in the analysis of factors associated with reporting, we must keep in mind that access to health care is likely to be an independent risk factor for underreporting in the general population. Another potential limitation is that the responses were dichotomous; there were substantial floor and ceiling effects resulting in failure to identify some of the exacerbations because once a symptom is present, no further change will be recorded.

Another limitation is that the study examines a relatively small prospective group of 116 patients monitored for only 6 months. However, as the population was enriched with patients with frequent exacerbators, 262 exacerbations were analyzed, which were equal to a rate of 2.25 per person every 6 months. This event rate can be explained because the relatively high proportion of patients included immediately after hospitalization might have contributed to a higher exacerbation rate and a possible seasonal effect. Missing data in the daily diary were also related to interest. The combination of both missing data and ceiling effects could have resulted in failure to identify some of the reported exacerbations in the daily diary. Furthermore, the analysis used ignored possible differences in symptom trajectory (early recovery from some symptoms) and these differences might be related to reporting.

Conclusions

This study evaluates compliance in the daily use of an mHealth app among patients with COPD having a documented history of exacerbations who are motivated in their self-care for long-term monitoring without a mentoring process or regular phone calls. The findings of this cohort study confirm that daily use of the Prevexair app is feasible and acceptable for reporting daily symptoms and medicine use among people with COPD who are motivated in their self-care because of frequent decompensations of their disease. In addition, this study shows that monitoring through the Prevexair app has the potential for implementation of self-care plans and offers opportunities for interventions in the treatment of patients at risk of frequent exacerbations, identifying symptoms and providing a better diagnosis of their chronic condition. Further research must be carried out to evaluate this strategy for the management of COPD in clinical practice. In the near future, mHealth apps will be a natural complement to health telematics and personal health records. They should be a part of a complete solution to address changes in health care provision, and they are particularly suitable for chronic disease prevention and management.

Acknowledgments

The authors thank the investigators, Francisco Javier Agustín Martínez, Pulmonology Department, Complejo Hospitalario U de Albacete, and Walther Ivan Giron Matute, Pulmonology Department, Hospital U Gregorio Marañón Madrid, Spain, who participated in the Prevexair study. The authors also thank Astra Zeneca for its financial support to carry out the study. This study has been promoted and sponsored by the Spanish Society of Pneumology and Thoracic Surgery. The financers had no role in study design, data collection, analysis, and decision to publish or in the preparation of this manuscript. This does not alter our adherence to the policies of the *Journal of Medical Internet Research* on sharing data and materials.

Authors' Contributions

JH, MCR, AG, LM, CD, FG, RR, and JS have intellectually contributed to this work and contributed to data analysis and interpretation of results. JH and MR wrote the manuscript. MF carried out the statistical analysis. All authors participated in drafting and revising the paper and assume accountability for all aspects of the work.

Conflicts of Interest

JH has received speaking fees from Boehringer Ingelheim and Gebro Pharma. This does not have a real or perceived conflict of interest between all these sources and this paper. MCR has received speaking fees from Boehringer Ingelheim, AstraZeneca, GlaxoSmithKline (GSK), Menarini, and Novartis and consulting fees from GSK, Gebro Pharma, and Novartis. This does not have a real or perceived conflict of interest between all these sources and this paper. LM has received travel coverings, funds for educative activities, research grants, paid advisories, and participated as Principal Investigator in randomized controlled trials sponsored by private companies or observational studies from Air Liquide, Astra Zeneca, Boston Scientific, Boeringher Ingelheim, Chiesi, ESTEVE, GSK Menarini, MSD, Novartis, Sanofi, Spanish Scientific Societies, and Government and the European Regional Cooperation Fund. This does not have a real or perceived conflict of interest between all these sources and this paper.
CD has received speaking fees from Menarini, GSK, Novartis, Chiesi, Teva, and Ferrer. This does not have a real or perceived conflict of interest between all these sources and this paper. FG has received speaking fees from GSK, Chiesi, Boehringer Ingelheim, Mundipharma, Menarini, Pfizer, Novartis, Esteve, Teva Pharmaceutical, Ferrer, Rovi, Astra Zeneca, Bial, and Actelion y Gebro Pharm. This does not have a real or perceived conflict of interest between all these sources and this paper.

References

1. Blasi F, Cesana G, Conti S, Chiodini V, Aliberti S, Fornari C, et al. The clinical and economic impact of exacerbations of chronic obstructive pulmonary disease: a cohort of hospitalized patients. PLoS One 2014;9(6):e101228 [FREE Full text] [doi: 10.1371/journal.pone.0101228] [Medline: 24971791]

2. Udsen FW, Lilholt PH, Hejlesen O, Ehlers L. Cost-effectiveness of telehealthcare to patients with chronic obstructive pulmonary disease: results from the Danish 'TeleCare North' cluster-randomised trial. BMJ Open 2017 May 17;7(5):e014616 [FREE Full text] [doi: 10.1136/bmjopen-2016-014616] [Medline: 28515193]
3. Wageck B, Cox NS, Holland AE. Recovery following acute exacerbations of chronic obstructive pulmonary disease - a review. Chronic Obstr Pulm Dis 2019 Feb;16(1):93-103. [doi: 10.1080/15412555.2019.1598965] [Medline: 31044644]
4. Seemungal TA, Wedzicha JA. Exacerbation frequency and FEV1 decline of COPD: is it geographic? Eur Respir J 2014 May;43(5):1220-1222 [FREE Full text] [doi: 10.1183/09031936.00046014] [Medline: 24789942]
5. Johansson G, Mushnikov V, Bäckström T, Engström A, Khalid JM, Wall J, et al. Exacerbations and healthcare resource utilization among COPD patients in a Swedish registry-based nation-wide study. BMC Pulm Med 2018 Jan 25;18(1):17 [FREE Full text] [doi: 10.1186/s12890-018-0573-0] [Medline: 29370846]
6. Esteban C, Castro-Acosta A, Alvarez-Martínez CJ, Capelastegui A, López-Campos JL, Pozo-Rodriguez F. Predictors of one-year mortality after hospitalization for an exacerbation of COPD. BMC Pulm Med 2018 Jan 25;18(1):18 [FREE Full text] [doi: 10.1186/s12890-018-0574-z] [Medline: 29370849]
7. Hartl S, Lopez-Campos JL, Pozo-Rodriguez F, Castro-Acosta A, Studnicka M, Kaiser B, et al. Risk of death and readmission of hospital-admitted COPD exacerbations: European COPD Audit. Eur Respir J 2016 Jan;47(1):113-121 [FREE Full text] [doi: 10.1183/13993003.01391-2014] [Medline: 26493806]
8. Singh D, Agusti A, Anzueto A, Barnes PJ, Bourbeau J, Celli BR, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease: the GOLD science committee report 2019. Eur Respir J 2019 May;53(5):pii: 1900164. [doi: 10.1183/13993003.00164-2019] [Medline: 30846476]
9. Miravitlles M, Soler-Cataluña JJ, Calle M, Molina J, Almagro P, Quintano JA, et al. Spanish Guidelines for Management of Chronic Obstructive Pulmonary Disease (GesEPOC) 2017. Pharmacological Treatment of Stable Phase. Arch Bronconeumol 2017 Jun;53(6):324-335 [FREE Full text] [doi: 10.1016/j.arbres.2017.03.018] [Medline: 28477954]
10. Rubio MC, Hermosa JL, Soler-Cataluña JJ, López-Campos JL, Navarrete BA, Soriano JB, y el Grupo Epoconsul. Medical care according to risk level and adaptation to Spanish COPD guidelines (Gesepoc): the Epoconsul study. Arch Bronconeumol 2018 May;54(5):270-279 [FREE Full text] [doi: 10.1016/j.arbres.2017.11.015] [Medline: 29361320]
11. Hurst JR, Vestbo J, Anzueto A, Locantore N, Müllerova H, Tal-Singer R, Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) Investigators. Susceptibility to exacerbation in chronic obstructive pulmonary disease. N Engl J Med 2010 Sep 16;363(12):1128-1138. [doi: 10.1056/NEJMoa0909883] [Medline: 20843247]
12. Quint JK, Donaldson GC, Hurst JR, Goldring JJP, Seemungal TR, Wedzicha JA. Predictive accuracy of patient-reported exacerbation frequency in COPD. Eur Respir J 2011 Mar;37(3):501-507 [FREE Full text] [doi: 10.1183/09031936.00035909] [Medline: 20650988]
13. Gregersen TL, Green A, Frausing E, Ringbæk T, Brøndum E, Ulrik CS. Do telemedical interventions improve quality of life in patients with COPD? A systematic review. Int J Chron Obstruct Pulmon Dis 2016;11:809-822 [FREE Full text] [doi: 10.2147/COPD.S96079] [Medline: 27143872]
14. Miravitlles M, D'Urzo A, Singh D, Koblizek V. Pharmacological strategies to reduce exacerbation risk in COPD: a narrative review. Respir Res 2016 Sep 10;17(1):112 [FREE Full text] [doi: 10.1186/s12931-016-0425-5] [Medline: 27613392]
15. Mackay AJ, Kostikas K, Murray L, Martinez FJ, Miravitlles M, Donaldson G, et al. Patient-reported outcomes for the detection, quantification, and evaluation of chronic obstructive pulmonary disease exacerbations. Am J Respir Crit Care Med 2018 Sep 15;198(6):730-738. [doi: 10.1164/rccm.201712-2482CI] [Medline: 29671615]
16. Calderazzo MA, Trujillo-Torralbo M, Finney LJ, Singanayagam A, Bakhsoliani E, Padmanaban V, et al. Inflammation and infections in unreported chronic obstructive pulmonary disease exacerbations. Int J Chron Obstruct Pulmon Dis 2019;14:823-832 [FREE Full text] [doi: 10.2147/COPD.S191946] [Medline: 31114182]
17. Langsetmo L, Platt RW, Ernst P, Bourbeau J. Underreporting exacerbation of chronic obstructive pulmonary disease in a longitudinal cohort. Am J Respir Crit Care Med 2008 Feb 15;177(4):396-401. [doi: 10.1164/rccm.200708-1290OC] [Medline: 18048806]
18. Boer LM, Bischoff EW, Borgijink X, Vercoulen JH, Akkermans RP, Kerstjens HA, et al. 'Exacerbation-free time' to assess the impact of exacerbations in patients with chronic obstructive pulmonary disease (COPD): a prospective observational study. NPJ Prim Care Respir Med 2018 Apr 3;28(1):12 [FREE Full text] [doi: 10.1038/s41533-018-0079-5] [Medline: 29615628]
19. Seemungal TA, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 1998 May;157(5 Pt 1):1418-1422. [doi: 10.1164/ajrccm.157.5.9709032] [Medline: 9603117]
20. Majothi S, Jolly K, Heneghan NR, Price MJ, Riley RD, Turner AM, et al. Supported self-management for patients with COPD who have recently been discharged from hospital: a systematic review and meta-analysis. Int J Chron Obstruct Pulmon Dis 2015;10:853-867 [FREE Full text] [doi: 10.2147/COPD.S74162] [Medline: 25995625]
21. Laue J, Melbye H, Risør MB. Self-treatment of acute exacerbations of chronic obstructive pulmonary disease requires more than symptom recognition - a qualitative study of COPD patients' perspectives on self-treatment. BMC Fam Pract 2017 Jan 25;18(1):8 [FREE Full text] [doi: 10.1186/s12875-017-0582-8] [Medline: 28122492]

22. Jahns RG. research2guidance. 500m People Will Be Using Healthcare Mobile Applications in 2015 URL: https://research2guidance.com/500m-people-will-be-using-healthcare-mobile-applications-in-2015-2/ [accessed 2020-01-20]
23. Gaveikaite V, Fischer C, Schonenberg H, Pauws S, Kitsiou S, Chouvarda I, et al. Telehealth for patients with chronic obstructive pulmonary disease (COPD): a systematic review and meta-analysis protocol. BMJ Open 2018 Sep 19;8(9):e021865 [FREE Full text] [doi: 10.1136/bmjopen-2018-021865] [Medline: 30232108]
24. Shah SA, Velardo C, Farmer A, Tarassenko L. Exacerbations in chronic obstructive pulmonary disease: identification and prediction using a digital health system. J Med Internet Res 2017 Mar 7;19(3):e69 [FREE Full text] [doi: 10.2196/jmir.7207] [Medline: 28270380]
25. Yañez AM, Guerrero D, de Alejo RP, Garcia-Rio F, Alvarez-Sala JL, Calle-Rubio M, et al. Monitoring breathing rate at home allows early identification of COPD exacerbations. Chest 2012 Dec;142(6):1524-1529. [doi: 10.1378/chest.11-2728] [Medline: 22797131]
26. Hennemann S, Beutel ME, Zwerenz R. Drivers and barriers to acceptance of web-based aftercare of patients in inpatient routine care: a cross-sectional survey. J Med Internet Res 2016 Dec 23;18(12):e337 [FREE Full text] [doi: 10.2196/jmir.6003] [Medline: 28011445]
27. Dou K, Yu P, Deng N, Liu F, Guan Y, Li Z, et al. Patients' acceptance of smartphone health technology for chronic disease management: a theoretical model and empirical test. JMIR Mhealth Uhealth 2017 Dec 6;5(12):e177 [FREE Full text] [doi: 10.2196/mhealth.7886] [Medline: 29212629]
28. Dumais KM, Dias N, Khurana L, Gary ST, Witherspoon B, Evans CJ, et al. Preferences for use and design of electronic patient-reported outcomes in patients with chronic obstructive pulmonary disease. Patient 2019 Dec;12(6):621-629. [doi: 10.1007/s40271-019-00376-9] [Medline: 31313271]
29. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to ehealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 6;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
30. Miravitlles M, Kruesmann F, Haverstock D, Perroncel R, Choudhri SH, Arvis P. Sputum colour and bacteria in chronic bronchitis exacerbations: a pooled analysis. Eur Respir J 2012 Jun;39(6):1354-1360 [FREE Full text] [doi: 10.1183/09031936.00042111] [Medline: 22034649]
31. Ancochea J, Badiola C, Duran-Tauleria E, Rio FG, Miravitlles M, Muñoz L, et al. [The EPI-SCAN survey to assess the prevalence of chronic obstructive pulmonary disease in Spanish 40-to-80-year-olds: protocol summary]. Arch Bronconeumol 2009 Jan;45(1):41-47 [FREE Full text] [doi: 10.1016/j.arbres.2008.06.001] [Medline: 19186298]
32. Soriano JB, Lamprecht B, Ramírez AS, Martinez-Camblor P, Kaiser B, Alfageme I, et al. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. Lancet Respir Med 2015 Jun;3(6):443-450. [doi: 10.1016/S2213-2600(15)00157-5] [Medline: 25995071]
33. Ministerio de Sanidad, Consumo y Bienestar Social. 2015. Informe Anual del Sistema Nacional de Salud 2015 URL: http://www.mscbs.gob.es/estadEstudios/estadisticas/sisInfSanSNS/tablasEstadisticas/Inf_Anual_SNS_2015.1.pdf [accessed 2020-01-20]
34. Ministerio de Sanidad, Consumo y Bienestar Social. Madrid: Ministerio de Sanidad y Politica Social; 2014. Actualización de la Estrategia en EPOC del Sistema Nacional de Salud: Informe enero de 2014 URL: http://www.mscbs.gob.es/organizacion/sns/planCalidadSNS/pdf/EPOC_version_junio_2014.pdf [accessed 2019-07-24]
35. Calle Rubio M, Casamor R, Miravitlles M. Identification and distribution of COPD phenotypes in clinical practice according to Spanish COPD Guidelines: the FENEPOC study. Int J Chron Obstruct Pulmon Dis 2017;12:2373-2383 [FREE Full text] [doi: 10.2147/COPD.S137872] [Medline: 28848338]
36. de Miguel-Diez J, Lopez-de-Andres A, Herandez-Barrera V, Jimenez-Trujillo I, Puente-Maestu L, Cerezo-Lajas A, et al. Effect of the economic crisis on the use of health and home care services among Spanish COPD patients. Int J Chron Obstruct Pulmon Dis 2018;13:725-739 [FREE Full text] [doi: 10.2147/COPD.S150308] [Medline: 29535513]
37. Rubio MC, López-Campos JL, Soler-Cataluña JJ, Navarrete BA, Soriano JB, González-Moro JM, EPOCONSUL Study. Variability in adherence to clinical practice guidelines and recommendations in COPD outpatients: a multi-level, cross-sectional analysis of the EPOCONSUL study. Respir Res 2017 Dec 2;18(1):200 [FREE Full text] [doi: 10.1186/s12931-017-0685-8] [Medline: 29197415]
38. 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]
39. Sobnath DD, Philip N, Kayyali R, Nabhani-Gebara S, Pierscionek B, Vaes AW, et al. Features of a mobile support app for patients with chronic obstructive pulmonary disease: literature review and current applications. JMIR Mhealth Uhealth 2017 Feb 20;5(2):e17 [FREE Full text] [doi: 10.2196/mhealth.4951] [Medline: 28219878]
40. Howcroft M, Walters EH, Wood-Baker R, Walters JA. Action plans with brief patient education for exacerbations in chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2016 Dec 19;12:CD005074 [FREE Full text] [doi: 10.1002/14651858.CD005074.pub4] [Medline: 27990628]
41. Lenferink A, Brusse-Keizer M, van der Valk PD, Frith PA, Zwerink M, Monninkhof EM, et al. Self-management interventions including action plans for exacerbations versus usual care in patients with chronic obstructive pulmonary disease. Cochrane

Database Syst Rev 2017 Aug 4;8:CD011682 [FREE Full text] [doi: 10.1002/14651858.CD011682.pub2] [Medline: 28777450]
42. Zwerink M, Brusse-Keizer M, van der Valk PD, Zielhuis GA, Monninkhof EM, van der Palen J, et al. Self management for patients with chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2014 Mar 19(3):CD002990. [doi: 10.1002/14651858.CD002990.pub3] [Medline: 24665053]
43. The Rt Hon Andrew Lansley CBE, Department of Health and Social Care. UK Government. 2012. GPs to 'prescribe' Apps for Patients URL: https://www.gov.uk/government/news/gps-to-prescribe-apps-for-patients [accessed 2020-01-20]
44. O'Reilly JF, Williams AE, Holt K, Rice L. Defining COPD exacerbations: impact on estimation of incidence and burden in primary care. Prim Care Respir J 2006 Dec;15(6):346-353 [FREE Full text] [doi: 10.1016/j.pcrj.2006.08.009] [Medline: 17064963]
45. Walters EH, Walters J, Wills KE, Robinson A, Wood-Baker R. Clinical diaries in COPD: compliance and utility in predicting acute exacerbations. Int J Chron Obstruct Pulmon Dis 2012;7:427-435 [FREE Full text] [doi: 10.2147/COPD.S32222] [Medline: 22848156]
46. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
47. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
48. Ding H, Karunanithi M, Ireland D, McCarthy L, Hakim R, Phillips K, et al. Evaluation of an innovative mobile health programme for the self-management of chronic obstructive pulmonary disease (MH-COPD): protocol of a randomised controlled trial. BMJ Open 2019 Apr 25;9(4):e025381 [FREE Full text] [doi: 10.1136/bmjopen-2018-025381] [Medline: 31028038]
49. Becker S, Brandl C, Meister S, Nagel E, Miron-Shatz T, Mitchell A, et al. Demographic and health related data of users of a mobile application to support drug adherence is associated with usage duration and intensity. PLoS One 2015;10(1):e0116980 [FREE Full text] [doi: 10.1371/journal.pone.0116980] [Medline: 25629939]
50. Zhao P, Yoo I, Lancey R, Varghese E. Mobile applications for pain management: an app analysis for clinical usage. BMC Med Inform Decis Mak 2019 May 30;19(1):106 [FREE Full text] [doi: 10.1186/s12911-019-0827-7] [Medline: 31146739]
51. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013 Apr 15;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
52. Willis KF, Robinson A, Wood-Baker R, Turner P, Walters EH. Participating in research: exploring participation and engagement in a study of self-management for people with chronic obstructive pulmonary disease. Qual Health Res 2011 Sep;21(9):1273-1282. [doi: 10.1177/1049732311405216] [Medline: 21474655]
53. van der Heijden M, Lucas PJ, Lijnse B, Heijdra YF, Schermer TR. An autonomous mobile system for the management of COPD. J Biomed Inform 2013 Jun;46(3):458-469 [FREE Full text] [doi: 10.1016/j.jbi.2013.03.003] [Medline: 23500485]
54. Ding B, Small M, Holmgren U. A cross-sectional survey of current treatment and symptom burden of patients with COPD consulting for routine care according to GOLD 2014 classifications. Int J Chron Obstruct Pulmon Dis 2017;12:1527-1537 [FREE Full text] [doi: 10.2147/COPD.S133793] [Medline: 28579771]
55. Miravitlles M, Ribera A. Understanding the impact of symptoms on the burden of COPD. Respir Res 2017 Apr 21;18(1):67 [FREE Full text] [doi: 10.1186/s12931-017-0548-3] [Medline: 28431503]
56. Celli B, Blasi F, Gaga M, Singh D, Vogelmeier C, Pegoraro V, et al. Perception of symptoms and quality of life - comparison of patients' and physicians' views in the COPD MIRROR study. Int J Chron Obstruct Pulmon Dis 2017;12:2189-2196 [FREE Full text] [doi: 10.2147/COPD.S136711] [Medline: 28794623]
57. Walentynowicz M, van Diest I, Raes F, van den Bergh O. Ways of encoding somatic information and their effects on retrospective symptom reporting. Br J Health Psychol 2017 May;22(2):362-378. [doi: 10.1111/bjhp.12234] [Medline: 28191712]

Abbreviations

AECOPD: acute exacerbations of COPD
CAT: Chronic Obstructive Pulmonary Disease Assessment Test
COPD: chronic obstructive pulmonary disease
FEV1: forced expiratory volume in one second
GSK: GlaxoSmithKline
mHealth: mobile health
mMRC: modified Medical Research Council

Edited by G Eysenbach; submitted 31.07.19; peer-reviewed by I Yang, C Stepnowsky, J Edwards; comments to author 08.10.19; revised version received 14.10.19; accepted 16.12.19; published 19.03.20.

Please cite as:
Rodriguez Hermosa JL, Fuster Gomila A, Puente Maestu L, Amado Diago CA, Callejas González FJ, Malo De Molina Ruiz R, Fuentes Ferrer ME, Álvarez Sala-Walther JL, Calle Rubio M
Compliance and Utility of a Smartphone App for the Detection of Exacerbations in Patients With Chronic Obstructive Pulmonary Disease: Cohort Study
JMIR Mhealth Uhealth 2020;8(3):e15699
URL: http://mhealth.jmir.org/2020/3/e15699/
doi:10.2196/15699
PMID:32191213

Original Paper

An Innovative Electronic Health Toolkit (Our Whole Lives for Chronic Pain) to Reduce Chronic Pain in Patients With Health Disparities: Open Clinical Trial

Paula Gardiner[1], MPH, MD; Salvatore D'Amico[2], BS; Man Luo[2], MPH; Niina Haas[3], MA

[1]Department of Family Medicine and Community Health, University of Massachusetts Medical School, Worcester, MA, United States

[2]Department of Family Medicine, Boston Medical Center, Boston, MA, United States

[3]BrightOutcome, Buffalo Grove, IL, United States

Corresponding Author:
Paula Gardiner, MPH, MD
Department of Family Medicine and Community Health
University of Massachusetts Medical School
55 Lake Street North
Worcester, MA, 01655
United States
Phone: 1 7744416716
Fax: 1 7744416212
Email: Paula.Gardiner@umassmed.edu

Abstract

Background: Chronic pain affects millions of Americans. Our Whole Lives, an electronic health (eHealth) toolkit for Chronic Pain (Our Whole Lives for Chronic Pain [OWLCP]), is a mind-body chronic pain management platform that teaches self-management strategies to reduce pain impact and pain medication use.

Objective: The primary goal of this study was to evaluate the feasibility of OWLCP in reducing pain impact and pain-related outcomes.

Methods: We conducted a pre-post clinical study (2 cohorts) to assess the feasibility of OWLCP usage among low-income patients with chronic pain. Outcome data, collected at baseline and 9 weeks, included Patient-Reported Outcomes Measurement Information System (PROMIS-29), pain self-efficacy, and pain medication use. In the statistical analysis, we used descriptive statistics, logistic regression, linear regression, and qualitative methods.

Results: Among the enrolled 43 participants, the average age was 50 years, (39/43) 91% were female, (16/43) 37% were black, and (7/43) 16% were Hispanic. From baseline to follow-up, the PROMIS measures showed a reduction in depression (P=.02), pain interference (P=.003), and average pain impact score (P=.007). Pain self-efficacy increased ((P<.001), whereas opioid use had a 13% reduction (P=.03).

Conclusions: The eHealth chronic pain management platform, OWLCP, is a potential tool to reduce the impact of chronic pain for low-income racially diverse populations.

(JMIR Mhealth Uhealth 2020;8(3):e14768) doi:10.2196/14768

KEYWORDS

minority health; internet; Web-based; complementary therapies; mindfulness-based stress reduction

Introduction

In the United States, more than half of all adults experience pain in any given year [1]. Chronic pain may be defined as pain that persists past normal tissue healing (>3 months) [2]. It is characterized by substantial suffering and associated with other comorbidities, such as insomnia, depression, fatigue, lowered mobility, and reduced quality of life [2]. Patients with chronic pain mostly receive care during hurried visits to the primary care provider (PCP) where they are prescribed pharmacological treatments (eg, opioids and medications) despite mixed evidence of their efficacy and increased risk of potentially dangerous side effects, including addiction and death [3-6]. Even when these treatments are effective in reducing pain, they may not improve mental and functional status and may actually increase

depression [7-9]. There is a need for easy access of evidence-based nonpharmacological treatment options to help patients with chronic pain.

The impact of chronic pain is particularly severe in populations with racial and socioeconomic disparities who receive less patient education, surgery, and specialty referrals [10-14]. Health disparities in chronic pain treatment substantially impact the patients' ability to work and function [12,15]. The reduced use of nonpharmacological options such as mindfulness-based interventions (MBIs; ie, meditation and yoga) by low-income patients is attributed to limited insurance coverage, therapies not being offered to them as a treatment option, structural barriers such as transportation, or lack of access to these options in their neighborhoods [16-19].

The internet and mobile technology is an accessible, convenient, and time-saving method to deliver health interventions [20,21]. Ziebland et al [22] demonstrated that people living with chronic pain are increasingly using the Web to find information, support, reassurance, encouragement, and practical advice for self-management. They also use the Web to compare experiences of treatment and offer advice and support to others. Chronic pain interventions delivered via technology are increasing because of factors such as acceptability to increase social connection, convenience for patients, and the ability to interact at one's own pace at home [23-26]. For example, internet delivery can make participation in a Web-based nonpharmacological skill acquisition intervention possible when it otherwise would not have been due to pain flares or reduced mobility [26]. In addition, the flexibility of internet delivery can be appealing for people who are busy managing appointments and treatments, those who want to bypass barriers related to cost and insurance coverage and time commitments for in-person treatment, and those who are reluctant to engage in in-person group interventions [27-29].

As of 2018, US technology trends indicate that 67% of people who earn less than US $30,000, 77% of Hispanic adults, and 75% of African Americans own a mobile phone [30]. Among the clinical literature on technology-delivered interventions, few studies exist on racially diverse and low-income patients with cancer, HIV, or obesity [31-34]. This also applies to racially diverse students and low-income patients with chronic pain [35]. Yet, mobile health is a promising area for health education and intervention delivery in health-disparate communities.

Development of Our Whole Lives 1.0

The intervention, Our Whole Lives (OWL 1.0), an eHealth Toolkit, was developed during a Patient Centered Outcomes Research Institute (PCORI) Contract AD 1304-6218/ClinicalTrials.gov ID NCT02262377 in 2014 to 2017 [36,37]. This randomized controlled trial (RCT) tested an Integrative Medical Group Visit (IMGV) care model for the usual care in low-income racially diverse patients with chronic pain and depression [36]. The IMGV incorporates key principles and practices of mindfulness adapted for patients who are racially and culturally diverse and have low health literacy levels [38]. The IMGV curriculum introduces patients to the fundamentals of evidence-based integrative medicine such as nutrition, lifestyle, stress reduction, exercise, and massage [39-41]. During the 21-week RCT study, participants had access to OWL as an adjunctive patient education website for the IMGV.

The user interface, visuals, videos, scripts, and resource pages for OWL 1.0 were created with input from a patient advisory group (PAG; patients with chronic pain and depression) and beta testing of IMGV socioeconomically and racially diverse patient cohorts (~20 patients) [42,43]. OWL's content was designed for patients with low health literacy (grades 5-8) and has been adapted for a diverse patient population by ensuring that the images of patients on the site are representative of the diverse and vulnerable population in the study, for example, visual images and pictures of patients from diverse racial and ethnicity backgrounds on the website. OWL mirrored the curriculum taught in an IMGV including mindfulness exercises and self-management home practices, interactive self-monitoring and self-directed learning, as well as social support through interactions on community blogs. It allowed the participants to access each session one at a time on a weekly basis. The participants were encouraged to participate by commenting on each video, audio, or other experiential activities and track their progress.

This pre-post clinical trial was conducted to test the second-generation OWL (version 2; Figure 1), which we will refer to as OWL for Chronic Pain (OWLCP) outside of an in-person IMGV to see if the application is feasible to use as a stand-alone intervention. The main outcomes were pain impact (such as pain severity, pain interference, and physical function) and pain-related outcomes (eg, depression, anxiety, fatigue, sleep disturbance, ability to participate in social roles and activities, pain self-efficacy, and pain medication use). Finally, to understand how OWLCP potentially changes behavior, we used the Health Education Impact Questionnaire (HEIQ) to look for changes in health-directed behavior, positive and active engagement in life, social integration and support, and emotional distress [44]. Using these outcomes, we evaluated pre-post effects and estimated effect sizes.

Figure 1. Screenshot of the Our Whole Lives for Chronic Pain (OWLCP) website.

Methods

Setting

This study was held at Boston Medical Center (BMC), a private, not-for-profit, academic medical center and the largest safety net hospital in New England. BMC is primarily funded by charities or the government. Approximately 70% of the patients come from underserved populations, such as low-income and older adults, who rely on government payors such as Medicaid, the Health Safety Net, and Medicare for their coverage; 32% do not speak English as a primary language.

Study Design

This prospective clinical trial enrolled participants with chronic pain between October 2016 and January 2018. We conducted two 9-week cohorts of approximately 40 patients with chronic pain.

Recruitment and Enrollment

Inclusion criteria were as follows: chronic pain ≥4 on a 0 to 10 pain scale for at least 12 weeks [34,35]; older than 18 years; and the ability to provide informed consent and understand website information in English. Exclusion criteria were as follows: a major medical event or another life event that would interfere with their ability to use the internet and participate in the intervention; not currently having access to the internet; not having an internet-enabled device to access the website; pregnant or planning to become pregnant in the next 3 months; and active substance use of alcohol, cocaine, or heroin.

Recruitment flyers were placed in BMC's primary care outpatient clinics, local Young Men's Christian Association (YMCA) and community organizations in the Boston area. Both men and women attend the YMCA, and it acts as a local community resource. Research assistants (RAs) reached out to PCPs, either by attending relevant provider meetings per department (ie, internal medicine and family medicine) or electronically notifying them of this study through an electronic medical record. When a PCP identified potentially eligible participants, they gave the participant a pamphlet about the study; thereafter interested participants contacted the study so they could be considered for the study. Study staff followed up with participants to determine interest and eligibility. Participants also could self-refer themselves to the study to be screened for study enrollment.

If an individual met the specified criteria, they were invited to meet in person at BMC with the RA. They were invited to review and sign the informed consent and collect baseline data. The RA provided clear detailed information about what the study involved. If the participant was unable to come to BMC in person, the consent and initial visit process was conducted over the phone with the RA. Verbal consent was given over the phone; however, informed consent was signed when the participant met with the RA in person.

Intervention Our Whole Lives 2.0

On the basis of the feedback from the RCT participants and the PAG, OWL, an electronic health toolkit, for Chronic Pain 2.0 (OWLCP) was developed in 2017. OWLCP is the version being tested during this feasibility trial. OWLCP is a password-protected internet-based platform stored on a server compliant to the Health Insurance Portability and Accountability Act; this server could be accessed with a tablet, computer, or

mobile phone. This website provided interactive self-monitoring (ie, pain, mood, and medication use), self-directed learning (ie, health topics, mindfulness, movement, and nutrition), and social support (ie, online community forum). Table 1 includes session names, home practice assignments (each ~20 min), themes, and activities. Changes from version 1 to version 2 included the following: revision of curriculum webpages from PDFs to interactive webpages, removing session 10, and adding a pain medication-use tracking tool.

Table 1. Our Whole Lives for Chronic Pain website curriculum.

Title of session	Home practice	Theme or activity
Online orientation	N/A[a]	Awareness of breath meditation, ground rules, introduction to mindfulness
Our reactions to stress	BS[b]	Nonpharmacological approaches to stress
Our bodies and healthy sleep	BS, M[c]	Nonpharmacological approaches to sleep
Movement and food as medicine	Alternate BS/CY[d]; M 6 of 7 days	Movement and healthy eating skills
Our bodies' response to pain	Alternate BS/CY; M 6 of 7 days	Nonpharmacological pain approaches to pain management
Our bodies and inflammation	Alternate BS/CY; M 6 of 7 days	Nonpharmacological approaches to treating inflammation
Our bodies and depression	Alternate BS/CY; loving kindness meditation 6 of 7 days	Nonpharmacological approaches to depression and challenging communications
Understanding the role of food in our body	Choice of BS, CY, M, or loving kindness meditation	Mindful eating
Wellness review	N/A	Wellness review

[a]N/A: not applicable.

[b]BS: body scan.

[c]M: meditation.

[d]CY: chair yoga.

OWLCP's functions include the following: (1) a daily measurement record (present mood, physical state of the body, and daily medication use); (2) a monitored community blog on which participants post their thoughts and respond to prompts for each session, and the blog is monitored daily by an RA or a clinician; and (3) home practice progress log where the participant may track what mind-body practices they completed (awareness of breath [AOB] meditation, sitting meditation, loving kindness meditation, chair yoga, and body scan). Each mind-body practice was recorded by a certified yoga or meditation teacher and approximately lasts 20 min (audio or video recordings).

OWLCP contains 10 videos that discuss health topics such as prevention and management of pain and associated conditions (such as stress reactivity, insomnia, poor nutrition, inflammation, and depression). Participants are taught to practice principles of mindfulness (AOB meditation, sitting meditation, loving kindness meditation, chair yoga, and body scan) at each session. Patients are encouraged to interact with OWLCP by commenting in an open text box after each video, audio, or other experiential activity to monitor their progress with home practice, such as their mood, pain, and pain medication use, and to choose resources relevant to them. Participants could review all or part of completed modules, earn puzzle pieces and checkmarks by completing audios and videos and selected tasks (eg, practices), and self-monitor (ie, view tables showing progress in mind-body activities, pain, and mood).

OWLCP's resources library (mind-body resources, low-cost recommendations for nonpharmacological treatments, poetry, community resources, and tips for health eating and recipes) provides a range of chronic pain self-management practices. The RA and primary investigator (PI) monitored the use of the platform, posted questions to facilitate conversations on the community page, and answered any relevant questions. In addition, the participants were given access to a private journal.

For this study, we held two in-person group orientations for participants on how to navigate the OWLCP website. During the orientation, a clinician (PG), assisted by an RA, demonstrated how to use the OWLCP system, log on, navigate through the sessions, complete self-assessments, and interact on the community blog page. Participants had continuous access to OWLCP for 9 weeks. After the orientation, RAs called all participants at weeks 1 and 4 to assess for adverse events, remind them to log on to the website, and check in about technology concerns or problems. Week 4 included a midpoint survey that measured satisfaction and the number of times the participant interacted with different website features (video, audios, webpages, and blog).

Data Collection

Demographics included the following: age, sex, race, ethnicity, primary language, education level, employment status, and yearly household income. Race was categorized into black/African American, white, or Other. Primary language was dichotomized into English and Non-English. Education level was categorized into high school/Generalized Education Development (GED) or less, some college or associates/no degree, and college graduate/postgraduate. Employment status

was categorized into working full-time/part-time, unemployed/retired, and sick leave/disability.

The following information was collected through self-reported questionnaires at baseline and 9 weeks: Patient-Reported Outcomes Measurement Information System (PROMIS-29) [45], HEIQ Version 2.0 [44], pain self-efficacy, Perceived Stress Scale (PSS), pain medication use, and Attitudes Toward Computers Questionnaire (ATCQ) [46].

The PROMIS-29, a 29-item measure, assesses 7 domains: anxiety, depression, fatigue, pain interference, physical function and sleep disturbance, the ability to participate in social roles and activities, and pain intensity [47]. Each domain is scored separately with options ranging in value from 1 to 5, except for the pain intensity item, which ranges in value from 0 to 10. The pain impact score ranges from 8 to 50, which is a summed score of physical function, pain interference, and pain intensity. Higher score means that the patient is more impacted by pain. PROMIS has been validated in low-income racially diverse patients [47]. HEIQ Version 2.0, an instrument for the comprehensive evaluation of patient education programs, is a 40-item survey separated into 8 domains: positive and active engagement in life, health-directed behavior, skill and technique acquisition, constructive attitudes and approaches, self-monitoring and insight, health services navigation, social integration and support, and emotional well-being. HEIQ uses a 5-point Likert-type scale ranging from *strongly disagree* (1) to *strongly agree* (5). Domain scores are calculated by adding the score of items within scales and dividing the sum by the number of items in a particular scale; therefore, all domain scores range between 1 and 4. Higher scores indicate higher levels of self-management ability, with the exception of emotional distress where higher scores indicate more distress [44].

Pain self-efficacy was measured with the Pain Self-Efficacy Scale. It is a sum of 10 items, each rated on a scale of 0 to 6. Higher scores indicate higher levels of confidence in self-managing pain [48]. PSS measures the degree to which situations in one's life are perceived as stressful in the past month. PSS is a sum of 4 items each with a 0 to 4 scale. Items 2 and 3 are reversely scored. Higher scores indicate higher levels of perceived stress [48].

Data for self-reported pain medication use in the past 7 days were recorded at baseline and 9 weeks. Medications were categorized as either opioids, nonsteroidal anti-inflammatory drugs (NSAIDs), or other medications. Opioid includes the following: MS-contin, vicodin, oxycodone, oxycontin, percocet, tramadol, tylenol with codeine #3, and other medications (suboxone, codeine, and methadone). NSAIDs include ibuprofen, naproxen, aspirin, and other (nabumetone, ketoprofen, celecoxib). Miscellaneous/other medication includes acetaminophen, cyclobenzaprine, gabapentin, and other medications (pregabalin, diazepam, Biofreeze, nortriptyline, lidocaine, naratriptan, Cymbalta, magnesium, tizanidine, baclofen, and valium).

ATCQ assesses seven dimensions of attitudes toward computers. We used two items of the seven dimensions: comfort and efficacy. All items are in a 5-point Likert scale format, with response options ranging from strongly disagree to strongly agree [46]. The OWLCP platform tracks the number of log-ins and minutes for all activities such as watching videos, body scans, chair yoga, and meditations. All blog entries were collected and categorized.

Surveys were administered either in person or on the phone or via an email invitation through REDcap (v9.1.0 Vanderbilt University)—a password-protected research tool. Participants received US $50.00 for their involvement with the cohort study. The initial US $25.00 was disbursed after completion of the 9-week survey, and the other US $25.00 was given after the participants attended a focus group. These funds were disbursed using BMC/s Clincard system.

Data Analysis

Descriptive statistics were used to analyze survey information and adverse events. Means and SD as well as frequencies and percentages were calculated for demographic characteristics. Means and SDs were also calculated for PROMIS-29, HEIQ, PSS, Pain Self Efficacy Scale, medication use, and ATCQ at baseline and 9 weeks. In terms of OWL usage data, we tracked and summed the average and total number of minutes of mind-body practice, the number of times participants blogged, the number of times participants used the journal, and the time spent on OWLCP website. We also summarized quotes from the blog.

For the PROMIS-29 questionnaire, the scores for 7 subscales, except for *pain intensity* subscale, were converted into standardized *t* scores and SDs and compared with a national distribution of standardized *t* scores and SDs (mean 50, SD 10). The 95% confidence intervals (CIs) were calculated for each subscale in PROMIS-29. Higher scores are associated with better outcomes for physical function and satisfaction of social role. For physical function, the scores ranged from 1 (least difficult) to 5 (most difficult). Therefore, we reverse-coded for questions for physical function to make it a positive subscale. The questions for satisfaction of social role were positively scored. Lower scores are associated with better outcomes for anxiety, depression, fatigue, sleep disturbance, and pain interference.

Finally, pain intensity was scored using a 0 to 10 scale. Means and SDs were calculated for this variable. A lower score was associated with a better outcome. Pain impact score ranged from 8 to 50, which is a summed score of physical function, pain interference, and pain intensity. A higher score meant being more impacted by pain [49].

In HEIQ, the effect sizes and changes in percentages were calculated based on the scoring instructions for HEIQ. Means for each subscale at baseline and follow-up were calculated, and the group-change effect size was calculated from the mean change.

To compare the results between baseline and follow-up, we applied *t* test and multivariate regressions. For continuous outcomes with normal distributions, we used the paired *t* test. For nonparametric continuous outcomes in PROMIS-29 and ATCQ, we applied longitudinal linear regressions with Poisson model and a time predictor to calculate the *P* values. For binary outcomes in medication use questionnaire, the longitudinal

logistic regression was used to compare baseline and follow-up by calculating the odds ratios (ORs) and CIs. We used multiple imputation method for missing data. For the blog posts from the OWLCP platform, the posts were categorized by themes. All quantitative analyses were conducted using SAS 9.3. (SAS Enterprise Miner 13.1, SAS Institute Inc). For the qualitative data collected, all blog posts were individually analyzed and coded by 2 RAs using modified grounded theory. The blog posts were independently coded, and a codebook was generated. The PI primarily served to resolve differences found between the 2 initial coders.

This study was approved by the BMC Institutional Review Board.

Results

The study flow, screening, and study enrollment are shown in Figure 2. For possible participation in this study, the study team contacted a total of 120 participants by phone who were either self-referred or referred by a clinician. Of the total, 66 participants agreed to be screened, 59 were eligible, and 43 were enrolled. Of these, 7 participants were screened and found ineligible for the study (Figure 2). After enrollment (n=43), 2 participants voluntarily withdrew from the study, and 5 participants were lost to follow-up. Of these, 36 participants completed follow-up data collection (17% did not complete study). Specifically for the two cohorts, of the 18 participants who started cohort 1, 17 completed, whereas of the 25 participants who started cohort 2, 19 completed.

Figure 2. CONSORT diagram.

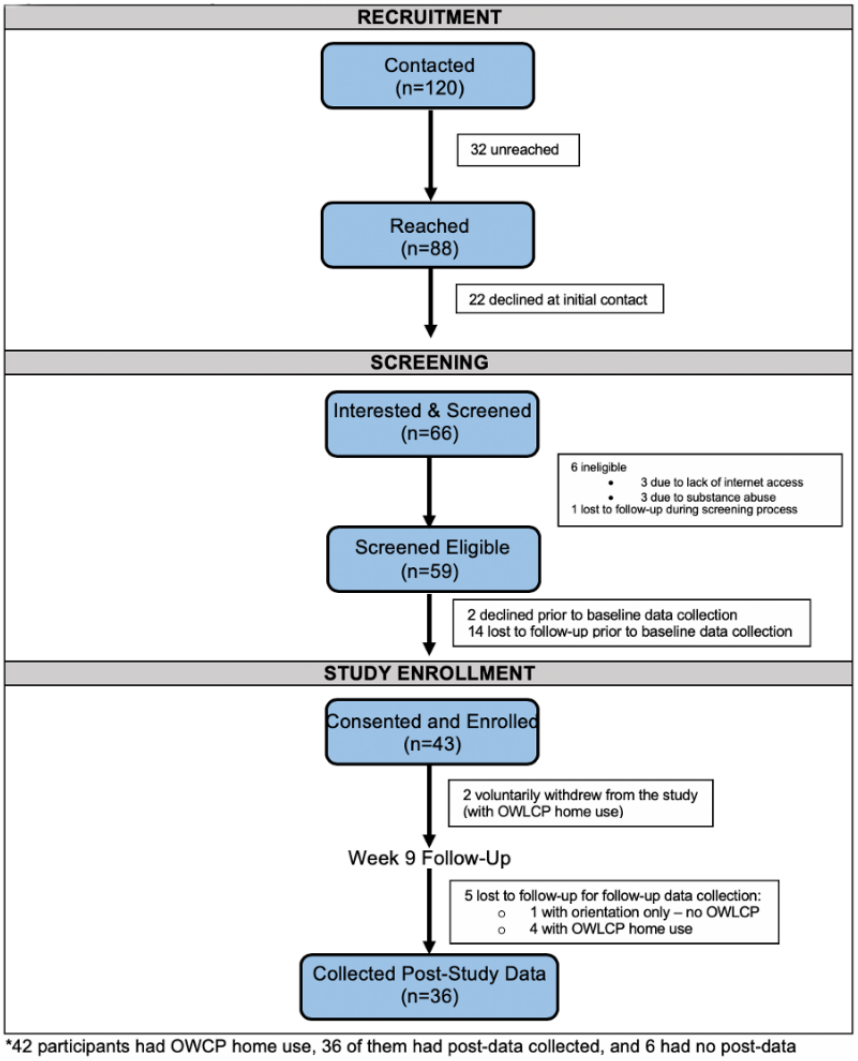


Table 2 lists all demographics factors. At baseline, the average age was 50 years, most participants were female (39/43, 91%), 37% (16/43) identified as black/African American, and 30% (13/43) identified as white. Of these, 23% of the participants (10/43) completed some high school. More than half of participants had some college degree or higher (33/43, 77%). Most participants were either unemployed/retired or on sick leave/disability (28/43, 65%). Approximately 26% (11/43) of the participants had a yearly household income of US $10,000 or less.

Table 2. Demographic characteristics.

Variables	Total (N=43)	Cohort 1 (n=18)	Cohort 2 (n=25)	*P* value
Age (years), mean (SD)	50.4 (12.6)	47.7 (11.9)	52.4 (12.9)	.23
Sex, n (%)				
Female	39 (91)	16 (89)	23 (92)	>.99
Race, n (%)				.46
Black/African American	16 (37)	8 (44)	8 (32)	
White	13 (30)	6 (33)	7 (28)	
Other[a]	14 (33)	4 (22)	10 (40)	
Hispanic/Latino, n (%)				
Yes	7 (16)	2 (11)	5 (20)	.68
Primary language, n (%)				
English	40 (93)	17 (94)	23 (92)	>.99
Education level, n (%)				.74
High school/generalized education development or less	10 (23)	5 (27)	5 (20)	
Some colleges or associates	16 (37)	7 (39)	9 (36)	
College or associate graduate, postgraduate	17 (40)	6 (33)	11 (44)	
Employment status, n (%)				.20
Working full time or part time	15 (35)	6 (33)	9 (36)	
Unemployed/retired/other[b]	11 (25)	7 (39)	4 (16)	
On sick leave/disability	17 (40)	5 (28)	12 (48)	
Yearly household income, n (%)				.22
US $10,000 or less	11 (26)	6 (33)	5 (20)	
US $10,001-US $90,000	16 (37)	4 (22)	12 (48)	
Refused/do not know	16 (37)	8 (44)	8 (32)	

[a]Other includes Native American (n=2), refused to answer (n=1), and other races (n=11).

[b]Other includes student (n=3) and other working status (n=3).

Table 3 lists the baseline and 9-week PROMIS-29 average t scores, means, SD, and 95% CI. At baseline, participants' physical function was 12 points lower than the national average standard t score (t score mean 38.2, SD 2.30). Table 3 shows a comparison of PROMIS-29 between baseline and follow-up. Depression decreased from baseline (t score mean 55.8, SD 2.87) to follow-up (t score mean 52.4, SD 3.40). This change was statistically significant (*P*=.02) with a large effect size (*d*=1.08). Satisfaction of social role increased from baseline (t score mean 40.1, SD 2.51) to follow-up (t score mean 42.9, SD 2.6; *P*=.09, d=1.10). Pain interference also showed a significant decrease (*P*=.003) with a large effect size (*d*=1.67) from baseline (t score mean 66.7, SD 2.20) to follow-up (t score mean 63.1, SD=2.10). Pain intensity decreased from baseline (mean 7.0, SD 1.48) to follow-up (mean 6.5, SD 2.22; *P*=.07, d=0.27). The decrease of pain impact was also statistically significant (*P*=.007) with a medium effect size (*d*=0.42) from baseline (mean 33.95, SD 7.4) to follow-up (mean 30.61, SD 8.53). There was no significant change from baseline to follow-up for physical function, anxiety, fatigue, and sleep disturbance.

Table 3. Patient-Reported Outcomes Measurement Information System-29 results: baseline and 9 weeks.

Item names	Baseline total (N=43)		9-week total (N=36)		*P* value	Effect size
	t score (*df*=42), mean (SD)	95% CI	*t* score (*df*=35), mean (SD)	95% CI		
Subscales (0-100)						
Physical function	38.2 (2.3)	33.6-42.7	38.0 (2.2)	33.7-42.4	.80	0.09
Anxiety	57.0 (3.4)	50.4-63.6	56.3 (3.4)	49.9-63.2	.98	0.21
Depression[a]	55.8 (2.9)	50.1-61.4	52.4 (3.4)	46.1-59.3	.02[b]	1.08[c]
Fatigue	59.9 (2.6)	54.9-64.9	57.7 (2.5)	53.0-62.9	.25	0.87[c]
Sleep disturbance	60.1 (3.6)	53.0-67.2	57.9 (3.5)	51.0-64.7	.19	0.62
Satisfaction of social role	40.1 (2.5)	35.2-45.0	42.9 (2.6)	37.7-48.1	.09	1.10[b]
Pain interference	66.7 (2.2)	62.3-71.0	63.1 (2.1)	58.9-67.2	.003[a]	1.67[b]
Subscales (0-10)						
Pain intensity	7.0 (1.5)	6.60-7.49	6.5 (2.2)	5.75-7.20	.07	0.27
Subscales (8-50)						
Pain impact	33.95 (7.4)	19.5-48.3	30.61 (8.5)	13.9-47.3	.007[a]	0.42

[a]Paired *t* test was used for depression, which was normally distributed. Regressions were applied to calculate *P* values for other subscales.

[b]The result is statistically significant at .05 level.

[c]The result is of a large effect size (Cohen d>0.8).

Table 4 shows the results of HEIQ among all the participants. There was a medium group-change effect size for skill and technique acquisition subscale (0.51). There was a small group-change effect size for health-directed behavior (0.35), positive and active engagement in life (0.28), and social integration and support (0.20). Net positive changes were seen for all domains. There was an increase in pain self-efficacy from baseline to 9 weeks (risk ratio [RR] 1.21 [95% CI 1.10-1.34]; *P*=.0001). No difference was seen in perceived stress score (RR 0.95 [95% CI 0.83-1.09]; *P*=.47).

Table 4. Health Education Impact Questionnaire results for all participants in Our Whole Lives for Chronic Pain study (N=35).

Subscale names	Baseline mean	Follow-up mean	Mean change	Group-change effect size[a]	Percent with reliable increase (%)	Percent with reliable decrease (%)	Net positive change (%)
Health-directed behavior	2.72	2.94	0.22	0.35	19	3	17
Positive and active engagement in life	2.99	3.17	0.15	0.28	19	6	14
Self-monitoring and insight	3.12	3.21	0.07	0.17	19	11	8
Constructive attitudes and approaches	3.09	3.16	0.04	0.08	11	8	3
Skill and technique acquisition	2.82	3.07	0.25	0.51	28	6	22
Social integration and support	2.84	2.98	0.12	0.20	19	6	14
Health services navigation	3.17	3.26	0.09	0.19	14	8	6
Emotional distress[b]	2.54	2.40	−0.10	−0.16	11	8	3

[a]Percentages are the proportions of participants who exceeded the threshold for reliable change.

[b]Percentages for emotional distress are reversed—the proportions in the positive reliable change cell are of those participants who had a reliably greater negative score on emotional distress at follow-up.

The medication use values of participants are presented in Table 5. At baseline, 74% (32/43) of all participants had any pain medication use in last 7 days, which increased to 83% (30/36) at 9 weeks (OR [95% CI]=1.69 [0.72-3.97], *P*=.23). Of these, 44% (19/43) used opioids at baseline, which decreased to 31% (11/36) at 9 weeks (OR [95% CI]=0.61 [0.39-0.94], *P*=.03). This was a statistically significant reduction in opioid use. In all, 51% (22/43) had used NSAIDs at baseline, which decreased to 44% (16/36) at 9 weeks (OR [95% CI]=0.73 [0.42-1.29], *P*=.28). There was no significant difference for other pain medication use between baseline (20/43, 47%) and 9 weeks (19/36, 53%) (OR [95% CI]=1.21 [0.60-2.46], *P*=.59).

Table 5. Pain medication use in the last week.

Pain medication	Baseline total (N=43)[a], n (%)	9-week total (N=36), n (%)
Medication use	32 (74)	30 (83)
Opioid use	19 (44)	11 (31)
Nonsteroidal anti-inflammatory drug use	22 (51)	16 (44)
Miscellaneous/other medication use	20 (47)	19 (53)

[a]A total of 11 participants did not use medication at baseline.

Our Whole Lives Use and Attitude Toward Computer Results

For results of participants' attitudes toward computers, two subscales—comfort and efficacy—were used. For the entire sample, there were no statistically significant increases for either comfort or efficacy (*P*=.53 and .57, respectively).

The time (in minutes) that participants spent on each activity from the OWLCP website across 9 weeks was noted. For a participant, the average number of minutes of use of OWLCP was 659 min (minimum=2, maximum=2352). The average numbers of days of use was 19 days (minimum=1 day, maximum=63 days). On average, participants spent a total of 61 min completing body scan, 45 min watching health topic videos, 25 min performing other meditation (AOB meditation, sitting meditation, or loving kindness meditation), and 24 min watching yoga videos. The mean number of log-ins per person was 25 (SD 24.9).

For the community blog, there were 348 posts involving 27 participants—average of 14 posts per person (minimum 0, maximum 51; 64% of the total sample), and for the private journal, there were 122 posts involving 27 participants (64% of the total sample; see Multimedia Appendix 1 for community board themes). No adverse events were reported.

Discussion

This is the first study to test an online clinician-monitored stand-alone self-management MBI system among urban diverse patients with chronic pain. We found a statistically significant increase in pain self-efficacy and a reduction in pain interference and depression. The average pain impact score decreased significantly, and the opioid use saw a 13% reduction. There were no significant changes from baseline to follow-up for physical function, anxiety, fatigue, and sleep disturbance. In terms of health education impact, there were increases in skill and technique acquisition, health-directed behavior, positive and active engagement in life, and social integration and support. Thus, we showed it is feasible for the OWLCP platform to function without an in-person medical group visit.

The findings of decreased pain interference, pain self-efficacy, and a reduction in average pain are consistent with those of studies of rural patients and international studies [50-54]. For example, Rini et al used an 8-week, automated, internet-based system called PainCoach, which included mind-body exercises in a racially diverse sample of participants with osteoarthritis in North Carolina. There was a significant pain reduction in women, and there were improvements in self-efficacy in both men and women [55,56]. In an Australian RCT of 148 participants with knee osteoarthritis, the PainCoach intervention significantly improved pain and physical function compared with the control group at 3 months. These improvements were sustained at 9 months [57]. Davis et al [51] studied an online 6-week MBI in patients with fibromyalgia, which showed an increase in self-efficacy for coping with pain.

Given the significant issues associated with chronic pain and its effect on work disability, social isolation, poor quality of life, and function, there is a great need for easily accessible culturally competent MBIs. An internet-based intervention that promotes nonpharmacological self-management and social support is an ideal approach in this population. Ziebland et al [22] noted that patients with chronic pain who made contact with others and learned from their experiences in managing pain emphasized the benefits of peer support. Furthermore, internet-delivered interventions would enhance accessibility for patients with chronic pain in rural areas that have difficulty with transportation or other physical limitations [58-61].

There have been several systematic reviews that have looked at Web-based technology for the treatment of chronic pain using nonpharmacological techniques such as MBIs [28,29,62-64]. In a systematic review of 16 studies by Toivonen et al [65], Web-based MBIs for people with chronic pain or fibromyalgia, irritable bowel syndrome, and other physical conditions showed positive effects compared with usual care on a variety of outcomes including pain acceptance, coping measures, and depressive symptoms as reported by most studies. In addition, systematic reviews have indicated that MBIs increase pain acceptance, pain tolerance, and ratings of life quality and reduce pain-associated psychological distress [66,67]. However, a more recent meta-analysis by Bawa et al [68] reviewed 11 studies that used only a randomized control group design provided less substantial effect sizes (eg, compared with control conditions) for clinical outcomes and considerable heterogeneity with regard to effect sizes. Furthermore, systematic reviews have assessed Web-supported MBIs on mental health reporting small to moderate beneficial effects of the interventions on depression, anxiety, stress, well-being, and mindfulness [69,70].

In total, 43% of our participants used opioids at baseline, which showed a statistically significant decrease from 44% to 31% at 9 weeks. However, there was an increase in any pain medication use, which increased from 74% to 83% at 9 weeks (OR [95% CI]=1.69 [0.72-3.97], *P*=.23). This was not statistically significant. This increase may be due to the increase in other types of pain medications such as acetaminophen, cyclobenzaprine, and gabapentin. For this analysis, we used

self-reported data, which may have been underestimated by participants. These findings need to be reproduced in a larger fully powered RCT.

A variety of factors may contribute to the benefits gained from using OWLCP as an MBI for patients with chronic pain. First, OWLCP was designed to simulate social connections by using an interactive blog. Using an iterative development process with PAGs and extensive beta testing, OWLCP was designed for low health literacy and diverse patient populations.

Many patients use the internet/mobile phone for social support, maintaining relationships with others, and finding health information [71]. In addition, the latest technology trends in the United States indicate that mobile phone adoption rates by those experiencing the highest rates of health disparities are increasing [31-34]. A recent systematic review documented significant improvements in outcome measures related to health behavior change using internet websites [72,73].

Limitations

Limitations to this feasibility study include a small sample size, short duration of the intervention, and self-reported information. By including only participants who had access to the internet, we may have biased the sample toward participants with favorable use of technology. This limitation could be addressed in future studies by providing a mobile phone and access to the internet. The OWLCP system was developed in English, thus excluding participants who were not fluent in English. Another limitation of the study is that it lacked a control group and had a larger sample of women compared with men. If this was a larger study, some of these limitations would be addressed. Although OWLCP recorded the number of minutes the participant was logged into the system, we do not know if the participant was actually practicing the mind-body techniques. We had no external sensors or video cameras to document objective measurement. On the flip side, we do not know if the participant was practicing outside of being logged into OWLCP.

Conclusions

In conclusion, we found strong evidence on the feasibility of OWLCP use by low-income, racially diverse patients with chronic pain as a stand-alone intervention. OWLCP increased pain-self efficacy and reduced pain interference and pain impact. We hypothesized that the benefits include increase in skill and technique acquisition, health-directed behavior, and social integration and support. However, future studies should focus on making OWLCP more accessible by removing the barrier of internet reliability (stand-alone application) and including other languages. Future studies should also add objective measurements for OWLCP.

Conflicts of Interest

The authors received funding from the Aetna Foundation.

Multimedia Appendix 1
A description of the qualitative data collected from the Community Board during the intervention.
[DOCX File , 15 KB - mhealth_v8i3e14768_app1.docx]

References

1. Nahin RL. Estimates of pain prevalence and severity in adults: United States, 2012. J Pain 2015 Aug;16(8):769-780 [FREE Full text] [doi: 10.1016/j.jpain.2015.05.002] [Medline: 26028573]
2. US Department of Health and Human Services. 2019. Pain Management Best Practices Inter-Agency Task Force Report URL: https://hhs.gov/sites/default/files/pmtf-final-report-2019-05-23.pdf [accessed 2019-11-18]
3. Chou R, Turner JA, Devine EB, Hansen RN, Sullivan SD, Blazina I, et al. The effectiveness and risks of long-term opioid therapy for chronic pain: a systematic review for a National Institutes of Health Pathways to Prevention Workshop. Ann Intern Med 2015 Feb 17;162(4):276-286. [doi: 10.7326/M14-2559] [Medline: 25581257]
4. Alford DP. Chronic back pain with possible prescription opioid misuse. J Am Med Assoc 2013 Mar 6;309(9):919-925. [doi: 10.1001/jama.2013.522] [Medline: 23462788]
5. Ray WA, Chung CP, Murray KT, Hall K, Stein CM. Prescription of long-acting opioids and mortality in patients with chronic noncancer pain. J Am Med Assoc 2016 Jun 14;315(22):2415-2423 [FREE Full text] [doi: 10.1001/jama.2016.7789] [Medline: 27299617]
6. Bohnert AS, Valenstein M, Bair MJ, Ganoczy D, McCarthy JF, Ilgen MA, et al. Association between opioid prescribing patterns and opioid overdose-related deaths. J Am Med Assoc 2011 Apr 6;305(13):1315-1321. [doi: 10.1001/jama.2011.370] [Medline: 21467284]
7. Martin BI, Deyo RA, Mirza SK, Turner JA, Comstock BA, Hollingworth W, et al. Expenditures and health status among adults with back and neck problems. J Am Med Assoc 2008 Feb 13;299(6):656-664. [doi: 10.1001/jama.299.6.656] [Medline: 18270354]
8. Scherrer JF, Salas J, Copeland LA, Stock EM, Ahmedani BK, Sullivan MD, et al. Prescription opioid duration, dose, and increased risk of depression in 3 large patient populations. Ann Fam Med 2016;14(1):54-62 [FREE Full text] [doi: 10.1370/afm.1885] [Medline: 26755784]

9. Scherrer JF, Svrakic DM, Freedland KE, Chrusciel T, Balasubramanian S, Bucholz KK, et al. Prescription opioid analgesics increase the risk of depression. J Gen Intern Med 2014 Mar;29(3):491-499 [FREE Full text] [doi: 10.1007/s11606-013-2648-1] [Medline: 24165926]
10. Simon LS. Relieving pain in America: a blueprint for transforming prevention, care, education, and research. J Pain Palliat Care Pharmacother 2012;26(2):197-198. [doi: 10.3109/15360288.2012.678473]
11. Newman AK, Van Dyke BP, Torres CA, Baxter JW, Eyer JC, Kapoor S, et al. The relationship of sociodemographic and psychological variables with chronic pain variables in a low-income population. Pain 2017 Sep;158(9):1687-1696. [doi: 10.1097/j.pain.0000000000000964] [Medline: 28570481]
12. Orhan C, van Looveren E, Cagnie B, Mukhtar NB, Lenoir D, Meeus M. Are pain beliefs, cognitions, and behaviors influenced by race, ethnicity, and culture in patients with chronic musculoskeletal pain: a systematic review. Pain Physician 2018 Nov;21(6):541-558 [FREE Full text] [doi: 10.36076/ppj.2018.6.541] [Medline: 30508984]
13. Hirsh AT, Hollingshead NA, Ashburn-Nardo L, Kroenke K. The interaction of patient race, provider bias, and clinical ambiguity on pain management decisions. J Pain 2015 Jun;16(6):558-568 [FREE Full text] [doi: 10.1016/j.jpain.2015.03.003] [Medline: 25828370]
14. Harle C, Apathy N, Cook R, Danielson EC, DiIulio J, Downs SM, et al. Information needs and requirements for decision support in primary care: an analysis of chronic pain care. AMIA Annu Symp Proc 2018;2018:527-534 [FREE Full text] [Medline: 30815093]
15. Janevic MR, McLaughlin SJ, Heapy AA, Thacker C, Piette JD. Racial and socioeconomic disparities in disabling chronic pain: findings from the health and retirement study. J Pain 2017 Dec;18(12):1459-1467 [FREE Full text] [doi: 10.1016/j.jpain.2017.07.005] [Medline: 28760648]
16. Escoto KH, Milbury K, Nguyen N, Cho D, Roberson C, Wetter D, et al. Use of complementary health practices in a church-based African American cohort. J Altern Complement Med 2018 Jun 8. [doi: 10.1089/acm.2018.0076] [Medline: 29883196]
17. Johnson C, Sheffield K, Brown R. Mind-body therapies for African-American women at risk for cardiometabolic disease: a systematic review. Evid Based Complement Alternat Med 2018;2018:5123217 [FREE Full text] [doi: 10.1155/2018/5123217] [Medline: 29681975]
18. Cheng T, D'Amico S, Luo M, Lestoquoy AS, Yinusa-Nyahkoon L, Laird LD, et al. Health disparities in access to nonpharmacologic therapies in an urban community. J Altern Complement Med 2019 Jan;25(1):48-60. [doi: 10.1089/acm.2018.0217] [Medline: 30234363]
19. Szanton SL, Wenzel J, Connolly AB, Piferi RL. Examining mindfulness-based stress reduction: perceptions from minority older adults residing in a low-income housing facility. BMC Complement Altern Med 2011 May 31;11:44 [FREE Full text] [doi: 10.1186/1472-6882-11-44] [Medline: 21627807]
20. Beatty L, Lambert S. A systematic review of internet-based self-help therapeutic interventions to improve distress and disease-control among adults with chronic health conditions. Clin Psychol Rev 2013 Jun;33(4):609-622. [doi: 10.1016/j.cpr.2013.03.004] [Medline: 23603521]
21. Rini C, Williams DA, Broderick JE, Keefe FJ. Meeting them where they are: using the internet to deliver behavioral medicine interventions for pain. Transl Behav Med 2012 Mar;2(1):82-92 [FREE Full text] [doi: 10.1007/s13142-011-0107-2] [Medline: 22924084]
22. Ziebland S, Lavie-Ajayi M, Lucius-Hoene G. The role of the internet for people with chronic pain: examples from the DIPEx International Project. Br J Pain 2015 Feb;9(1):62-64 [FREE Full text] [doi: 10.1177/2049463714555438] [Medline: 26516557]
23. Andersson G, Titov N. Advantages and limitations of internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
24. Cuijpers P, Marks IM, van Straten A, Cavanagh K, Gega L, Andersson G. Computer-aided psychotherapy for anxiety disorders: a meta-analytic review. Cogn Behav Ther 2009;38(2):66-82. [doi: 10.1080/16506070802694776] [Medline: 20183688]
25. 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 Nov 1;1(2):e24 [FREE Full text] [doi: 10.2196/mhealth.2733] [Medline: 25099314]
26. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. J Med Internet Res 2017 Mar 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
27. Wahbeh H, Svalina MN, Oken BS. Group, one-on-one, or internet? Preferences for mindfulness meditation delivery format and their predictors. Open Med J 2014;1:66-74 [FREE Full text] [doi: 10.2174/1874220301401010066] [Medline: 27057260]
28. Macea DD, Gajos K, Calil YA, Fregni F. The efficacy of web-based cognitive behavioral interventions for chronic pain: a systematic review and meta-analysis. J Pain 2010 Oct;11(10):917-929. [doi: 10.1016/j.jpain.2010.06.005] [Medline: 20650691]
29. Ashraf EM. Internet-based interventions for pain management: a systematic review of randomised controlled trial (RCTs) conducted from 2010 to 2014. J Public Health Epidemiol 2015;7(5):170-182. [doi: 10.5897/jphe2014.0702]

30. Pew Research Center. 2019 Jun 12. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-11-18]
31. Anderson-Lewis C, Darville G, Mercado RE, Howell S, di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the United States: systematic literature review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
32. Billings DW, Leaf SL, Spencer J, Crenshaw T, Brockington S, Dalal RS. A randomized trial to evaluate the efficacy of a web-based HIV behavioral intervention for high-risk African American women. AIDS Behav 2015 Jul;19(7):1263-1274 [FREE Full text] [doi: 10.1007/s10461-015-0999-9] [Medline: 25616838]
33. Joseph RP, Dutton GR, Cherrington A, Fontaine K, Baskin M, Casazza K, et al. Feasibility, acceptability, and characteristics associated with adherence and completion of a culturally relevant internet-enhanced physical activity pilot intervention for overweight and obese young adult African American women enrolled in college. BMC Res Notes 2015 Jun 2;8:209. [doi: 10.1186/s13104-015-1159-z] [Medline: 26032016]
34. Staffileno BA, Tangney CC, Fogg L. Favorable outcomes using an eHealth approach to promote physical activity and nutrition among young African American women. J Cardiovasc Nurs 2018;33(1):62-71. [doi: 10.1097/JCN.0000000000000409] [Medline: 28481823]
35. Berman RL, Iris MA, Bode R, Drengenberg C. The effectiveness of an online mind-body intervention for older adults with chronic pain. J Pain 2009 Jan;10(1):68-79. [doi: 10.1016/j.jpain.2008.07.006] [Medline: 18774342]
36. Gardiner P, Lestoquoy AS, Gergen-Barnett K, Penti B, White LF, Saper R, et al. Design of the integrative medical group visits randomized control trial for underserved patients with chronic pain and depression. Contemp Clin Trials 2017 Mar;54:25-35. [doi: 10.1016/j.cct.2016.12.013] [Medline: 27979754]
37. Shamekhi A, Bickmore T, Lestoquoy A, Negash L, Gardiner P. Blissful agents: adjuncts to group medical visits for chronic pain and depression. In: Intelligent Virtual Agents: 16th International Conference. Cham: Springer; 2016.
38. Cornelio-Flores O, Lestoquoy AS, Abdallah S, DeLoureiro A, Lorente K, Pardo B, et al. The Latino integrative medical group visit as a model for pain reduction in underserved Spanish speakers. J Altern Complement Med 2018 Feb;24(2):125-131. [doi: 10.1089/acm.2017.0132] [Medline: 28885858]
39. Lee C, Crawford C, Swann S, Active Self-Care Therapies for Pain (PACT) Working Group. Multimodal, integrative therapies for the self-management of chronic pain symptoms. Pain Med 2014 Apr;15(Suppl 1):S76-S85. [doi: 10.1111/pme.12408] [Medline: 24734863]
40. Delgado R, York A, Lee C, Crawford C, Buckenmaier C, Schoomaker E, Active Self-Care Therapies for Pain (PACT) Working Group. Assessing the quality, efficacy, and effectiveness of the current evidence base of active self-care complementary and integrative medicine therapies for the management of chronic pain: a rapid evidence assessment of the literature. Pain Med 2014 Apr;15(Suppl 1):S9-20. [doi: 10.1111/pme.12412] [Medline: 24734865]
41. Goode AP, Coeytaux RR, McDuffie J, Duan-Porter W, Sharma P, Mennella H, et al. An evidence map of yoga for low back pain. Complement Ther Med 2016 Apr;25:170-177. [doi: 10.1016/j.ctim.2016.02.016] [Medline: 27062965]
42. Dresner D, Resnick K, Gardiner P, Barnett KG, Laird L. Qualitative evaluation of an integrative medicine group visits program for patients with chronic pain and associated comorbidities. J Altern Complement Med 2014;20(5):A55-A56. [doi: 10.1089/acm.2014.5144.abstract]
43. Gardiner P, Crooks D, Shamekhi A, McCue K, Bickmore TW. Patient Advisory Groups and Integrative Medicine: How and Why to Incorporate them into Research Studies. In: Proceedings of the IM4US 5th Annual Conference for Integrative Medicine. 2015 Presented at: IM4US'15; August 6-8, 2015; Boston, MA.
44. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ Couns 2007 May;66(2):192-201. [doi: 10.1016/j.pec.2006.12.002] [Medline: 17320338]
45. Deyo RA, Ramsey K, Buckley DI, Michaels L, Kobus A, Eckstrom E, et al. Performance of a Patient Reported Outcomes Measurement Information System (PROMIS) short form in older adults with chronic musculoskeletal pain. Pain Med 2016 Feb;17(2):314-324 [FREE Full text] [doi: 10.1093/pm/pnv046] [Medline: 26814279]
46. Reznikoff M, Holland CH, Stroebel CF. Attitudes toward computers among employees of a psychiatric hospital. Ment Hyg 1967 Jul;51(3):419-425. [doi: 10.1097/00006199-196801000-00052] [Medline: 6057531]
47. Chen WH, Revicki DA, Amtmann D, Jensen MP, Keefe FJ, Cella D. Development and analysis of PROMIS pain intensity scale. Qual Life Res 2012;20(Suppl 1):18 [FREE Full text]
48. Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. Eur J Pain 2007 Feb;11(2):153-163. [doi: 10.1016/j.ejpain.2005.12.008] [Medline: 16446108]
49. Amtmann D, Cook KF, Jensen MP, Chen W, Choi S, Revicki D, et al. Development of a PROMIS item bank to measure pain interference. Pain 2010 Jul;150(1):173-182 [FREE Full text] [doi: 10.1016/j.pain.2010.04.025] [Medline: 20554116]
50. Gardner-Nix J, Backman S, Barbati J, Grummitt J. Evaluating distance education of a mindfulness-based meditation programme for chronic pain management. J Telemed Telecare 2008;14(2):88-92. [doi: 10.1258/jtt.2007.070811] [Medline: 18348755]

51. Davis MC, Zautra AJ. An online mindfulness intervention targeting socioemotional regulation in fibromyalgia: results of a randomized controlled trial. Ann Behav Med 2013 Dec;46(3):273-284. [doi: 10.1007/s12160-013-9513-7] [Medline: 23670111]
52. 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 Jan 9;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]
53. Zernicke KA, Campbell TS, Speca M, McCabe-Ruff K, Flowers S, Carlson LE. A randomized wait-list controlled trial of feasibility and efficacy of an online mindfulness-based cancer recovery program: the eTherapy for cancer applying mindfulness trial. Psychosom Med 2014 May;76(4):257-267. [doi: 10.1097/PSY.0000000000000053] [Medline: 24804884]
54. Dowd H, Hogan MJ, McGuire BE, Davis MC, Sarma KM, Fish RA, et al. Comparison of an online mindfulness-based cognitive therapy intervention with online pain management psychoeducation: a randomized controlled study. Clin J Pain 2015 Jun;31(6):517-527. [doi: 10.1097/AJP.0000000000000201] [Medline: 25565584]
55. Rini C, Porter LS, Somers TJ, McKee DC, Keefe FJ. Retaining critical therapeutic elements of behavioral interventions translated for delivery via the internet: recommendations and an example using pain coping skills training. J Med Internet Res 2014 Dec 19;16(12):e245 [FREE Full text] [doi: 10.2196/jmir.3374] [Medline: 25532216]
56. Rini C, Porter LS, Somers TJ, McKee DC, DeVellis RF, Smith M, et al. Automated internet-based pain coping skills training to manage osteoarthritis pain: a randomized controlled trial. Pain 2015 May;156(5):837-848 [FREE Full text] [doi: 10.1097/j.pain.0000000000000121] [Medline: 25734997]
57. Bennell KL, Nelligan R, Dobson F, Rini C, Keefe F, Kasza J, et al. Effectiveness of an internet-delivered exercise and pain-coping skills training intervention for persons with chronic knee pain: a randomized trial. Ann Intern Med 2017 Apr 4;166(7):453-462. [doi: 10.7326/M16-1714] [Medline: 28241215]
58. Winzelberg AJ, Classen C, Alpers GW, Roberts H, Koopman C, Adams RE, et al. Evaluation of an internet support group for women with primary breast cancer. Cancer 2003 Mar 1;97(5):1164-1173 [FREE Full text] [doi: 10.1002/cncr.11174] [Medline: 12599221]
59. Yanez B, McGinty HL, Mohr DC, Begale MJ, Dahn JR, Flury SC, et al. Feasibility, acceptability, and preliminary efficacy of a technology-assisted psychosocial intervention for racially diverse men with advanced prostate cancer. Cancer 2015 Dec 15;121(24):4407-4415 [FREE Full text] [doi: 10.1002/cncr.29658] [Medline: 26348661]
60. Penedo FJ, Dahn JR, Molton I, Gonzalez JS, Kinsinger D, Roos BA, et al. Cognitive-behavioral stress management improves stress-management skills and quality of life in men recovering from treatment of prostate carcinoma. Cancer 2004 Jan 1;100(1):192-200 [FREE Full text] [doi: 10.1002/cncr.11894] [Medline: 14692040]
61. Duffecy J, Sanford S, Wagner L, Begale M, Nawacki E, Mohr DC. Project onward: an innovative e-health intervention for cancer survivors. Psychooncology 2013 Apr;22(4):947-951 [FREE Full text] [doi: 10.1002/pon.3075] [Medline: 22438297]
62. Garg S, Garg D, Turin TC, Chowdhury MFU. 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]
63. Bender JL, Radhakrishnan A, Diorio C, Englesakis M, Jadad AR. Can pain be managed through the internet? A systematic review of randomized controlled trials. Pain 2011 Aug;152(8):1740-1750. [doi: 10.1016/j.pain.2011.02.012] [Medline: 21565446]
64. Heapy AA, Higgins DM, Cervone D, Wandner L, Fenton BT, Kerns RD. A systematic review of technology-assisted self-management interventions for chronic pain: looking across treatment modalities. Clin J Pain 2015 Jun;31(6):470-492. [doi: 10.1097/AJP.0000000000000185] [Medline: 25411862]
65. Toivonen KI, Zernicke K, Carlson LE. Web-based mindfulness interventions for people with physical health conditions: systematic review. J Med Internet Res 2017 Aug 31;19(8):e303 [FREE Full text] [doi: 10.2196/jmir.7487] [Medline: 28860106]
66. Chiesa A, Serretti A. Mindfulness-based interventions for chronic pain: a systematic review of the evidence. J Altern Complement Med 2011 Jan;17(1):83-93. [doi: 10.1089/acm.2009.0546] [Medline: 21265650]
67. Veehof MM, Oskam M, Schreurs KM, Bohlmeijer ET. Acceptance-based interventions for the treatment of chronic pain: a systematic review and meta-analysis. Pain 2011 Mar;152(3):533-542 [FREE Full text] [doi: 10.1016/j.pain.2010.11.002] [Medline: 21251756]
68. Bawa FL, Mercer SW, Atherton RJ, Clague F, Keen A, Scott NW, et al. Does mindfulness improve outcomes in patients with chronic pain? Systematic review and meta-analysis. Br J Gen Pract 2015 Jun;65(635):e387-e400 [FREE Full text] [doi: 10.3399/bjgp15X685297] [Medline: 26009534]
69. Spijkerman M, Pots W, Bohlmeijer E. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114 [FREE Full text] [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
70. Zhao D, Lustria ML, Hendrickse J. Systematic review of the information and communication technology features of web- and mobile-based psychoeducational interventions for depression. Patient Educ Couns 2017 Jun;100(6):1049-1072. [doi: 10.1016/j.pec.2017.01.004] [Medline: 28126383]
71. Ziebland S, Wyke S. Health and illness in a connected world: how might sharing experiences on the internet affect people's health? Milbank Q 2012 Jun;90(2):219-249 [FREE Full text] [doi: 10.1111/j.1468-0009.2012.00662.x] [Medline: 22709387]

72. Vance K, Howe W, Dellavalle RP. Social internet sites as a source of public health information. Dermatol Clin 2009 Apr;27(2):133-6, vi. [doi: 10.1016/j.det.2008.11.010] [Medline: 19254656]
73. Maher CA, Lewis LK, Ferrar K, Marshall S, de Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]

Abbreviations

AOB: awareness of breath
ATCQ: Attitudes Toward Computers Questionnaire
BMC: Boston Medical Center
HEIQ: Health Education Impact Questionnaire
IMGV: Integrative Medical Group Visit
MBI: mindfulness-based interventions
NSAIDs: nonsteroidal anti-inflammatory drugs
OR: odds ratio
OWL: Our Whole Lives
OWLCP: OWL for Chronic Pain
PAG: patient advisory group
PCP: primary care provider
PI: primary investigator
PROMIS-29: Patient-Reported Outcomes Measurement Information System
PSS: Perceived Stress Scale
RA: research assistant
RCT: randomized controlled trial
RR: risk ratio
YMCA: Young Men's Christian Association

Edited by G Eysenbach; submitted 20.05.19; peer-reviewed by C Greco, I Truccolo, M Steen, J Mitchell; comments to author 28.06.19; revised version received 29.08.19; accepted 22.10.19; published 30.03.20.

Please cite as:
Gardiner P, D'Amico S, Luo M, Haas N
An Innovative Electronic Health Toolkit (Our Whole Lives for Chronic Pain) to Reduce Chronic Pain in Patients With Health Disparities: Open Clinical Trial
JMIR Mhealth Uhealth 2020;8(3):e14768
URL: https://mhealth.jmir.org/2020/3/e14768
doi:10.2196/14768
PMID:32224487

Original Paper

A Culturally Relevant Smartphone-Delivered Physical Activity Intervention for African American Women: Development and Initial Usability Tests of Smart Walk

Rodney P Joseph[1], PhD; Colleen Keller[1], PhD; Sonia Vega-López[2,3], PhD; Marc A Adams[2], MPH, PhD; Rebekah English[1], BS; Kevin Hollingshead[2], BSEET; Steven P Hooker[4], PhD; Michael Todd[5], PhD; Glenn A Gaesser[2], PhD; Barbara E Ainsworth[2,6], MPH, PhD

[1]Center for Health Promotion and Disease Prevention, Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

[2]College of Health Solutions, Arizona State University, Phoenix, AZ, United States

[3]Southwest Interdisciplinary Research Center, Arizona State University, Phoenix, AZ, United States

[4]College of Health and Human Services, San Diego State University, San Diego, CA, United States

[5]Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

[6]Shanghai University of Sport, Yangpu Qu, Shanghai Shi, China

Corresponding Author:
Rodney P Joseph, PhD
Center for Health Promotion and Disease Prevention
Edson College of Nursing and Health Innovation
Arizona State University
550 N 3rd St
Phoenix, AZ, 85004
United States
Phone: 1 6024960772
Email: rodney.joseph@asu.edu

Abstract

Background: *Smart Walk* is a culturally relevant, social cognitive theory–based, smartphone-delivered intervention designed to increase physical activity (PA) and reduce cardiometabolic disease risk among African American (AA) women.

Objective: This study aimed to describe the development and initial usability testing results of *Smart Walk*.

Methods: *Smart Walk* was developed in 5 phases. Phases 1 to 3 focused on initial intervention development, phase 4 involved usability testing, and phase 5 included intervention refinement based on usability testing results. In phase 1, a series of 9 focus groups with 25 AA women (mean age 38.5 years, SD 7.8; mean BMI 39.4 kg/m2, SD 7.3) was used to identify cultural factors associated with PA and ascertain how constructs of social cognitive theory can be leveraged in the design of a PA intervention. Phase 2 included the analysis of phase 1 qualitative data and development of the structured PA intervention. Phase 3 focused on the technical development of the smartphone app used to deliver the intervention. Phase 4 consisted of a 1-month usability trial of *Smart Walk* (n=12 women; mean age 35.0 years, SD 8.5; mean BMI 40 kg/m2, SD 5.0). Phase 5 included refinement of the intervention based on the usability trial results.

Results: The 5-phase process resulted in the development of the *Smart Walk* smartphone-delivered PA intervention. This PA intervention was designed to target social cognitive theory constructs of behavioral capability, outcome expectations, social support, self-efficacy, and self-regulation and address deep structure sociocultural characteristics of collectivism, racial pride, and body appearance preferences of AA women. Key features of the smartphone app included (1) personal profile pages, (2) multimedia PA promotion modules (ie, electronic text and videos), (3) discussion boards, and (4) a PA self-monitoring tool. Participants also received 3 PA promotion text messages each week.

Conclusions: The development process of *Smart Walk* was designed to maximize the usability, cultural relevance, and impact of the smartphone-delivered PA intervention.

(JMIR Mhealth Uhealth 2020;8(3):e15346) doi:10.2196/15346

KEYWORDS

eHealth; mHealth; exercise; minority health; primary prevention; heart diseases; African-American

Introduction

Background

Regular physical activity (PA) is an independent risk factor for the prevention and control of cardiometabolic diseases, including obesity, cardiovascular disease, and type 2 diabetes [1-3]. Despite the health benefits of PA, many African American (AA) women are insufficiently active (ie, do not achieve the US aerobic PA guidelines of 150 min per week of moderate-intensity PA, 75 min of vigorous PA, or an equivalent combination of durations and intensities [4]). Recent estimates from 3 national surveys measuring population-based PA indicate only 27% to 40% of AA women achieve the goals stated in national aerobic PA guidelines [5]. These low PA levels likely contribute to a high prevalence of obesity (56%) [6], cardiovascular disease (48%) [7], and type 2 diabetes (13%) [8] in AA women [9].

Emerging evidence from the proliferation of electronic health (eHealth) and mobile health (mHealth) PA interventions largely supports the preliminary efficacy of such interventions to promote PA [10-14] and improve various cardiometabolic disease risk factors [14]. A 2017 meta-analysis [15] highlighted the effectiveness of eHealth and mHealth PA interventions when compared with non–technology-delivered PA interventions. The authors examined randomized trials testing technology-delivered PA interventions (ie, interventions delivered using mobile phones, websites, social media, and email) vs non–technology-delivered PA interventions. The results indicated that technology-delivered PA interventions were 12% more effective for increasing PA than non–technology-delivered PA intervention arms. A more recent 2019 meta-analysis [16] of smartphone-delivered PA interventions further supported the use of innovative technologies to deliver PA interventions, with results showing smartphone-delivered interventions were efficient in increasing both minutes (ie, 10.5 min) and steps per day (ie, 735 steps) of PA.

eHealth and mHealth PA interventions also provide several advantages when compared with traditional face-to-face methods of intervention delivery. From a participant's perspective, participants have the ability to access PA intervention materials virtually anywhere and at times convenient with their daily schedules. These features can help overcome barriers associated with in-person intervention delivery, including transportation issues and balancing work and family life schedules to attend intervention sessions [17-19]. From a research standpoint, technology-delivered interventions provide the opportunity for researchers to optimize theoretical fidelity (ie, defined by Rovniak et al [20] as the precision in replicating theory-based recommendations) of the behavior change principles underpinning their interventions. For example, PA interventions often use self-monitoring and feedback as methods to increase self-efficacy for PA (ie, targeting self-efficacy sources of mastery experiences and verbal persuasion). Before the development of eHealth and mHealth technologies, many researchers relied on participants to wear pedometers and manually record their PA levels in diaries. These diaries were then provided to the research team at a given time interval for review (ie, once a week or once a month) and to formulate feedback on participant progress. Current technology allows this process to occur in real time with less participant burden through the use of commercially available activity monitors, which can be viewed as a closer match to theoretical ideals. Similarly, social support for PA can be facilitated through asynchronous Web-based discussion boards, text messages, and video chats, as opposed to the traditional in-person social support sessions held at specified locations at predetermined dates and times.

Despite the favorable PA outcomes and advantages of using eHealth and mHealth approaches to deliver PA interventions, limited research has explored the efficacy of these types of interventions among AA women. This observation is surprising, given AA women engage in significantly lower PA levels than men and women of other race/ethnicities [21,22] and use electronic and mobile communication technologies (social media, internet, and mobile/smartphones) at equal or greater levels [23,24]. A recent review of eHealth and mHealth PA interventions [25] identified only 6 studies focused on AA women. Among these studies, all were preliminary in nature, and none included a smartphone-based app to promote PA (ie, 3 studies used websites, 2 used text messages, and 1 used the social media website Facebook and text messages). PA outcomes of these studies were generally favorable, with 4 (67%) reporting positive outcomes for at least one PA outcome measure. However, because of heterogeneity of study designs, multicomponent nature of interventions evaluated, and PA assessment methods, the authors were unable to determine if 1 method of intervention delivery was more effective than the another. Another main finding of this review was the lack of cultural tailoring efforts employed by researchers. Only 3 studies focusing on AA women (ie, 50%) reported some type of cultural tailoring—2 were tailored at the surface level and 1 at the deep structure level. Surface-level cultural tailoring involves matching the packaging of a health promotion intervention to the overt social and behavioral characteristics of the intended population (ie, including images of AA women and statistics regarding the PA health disparities of AA women [26]). *Deep structure* cultural tailoring involves acknowledging a group's sociocultural values, beliefs, and behaviors and harnessing these phenomena to promote behavior change [26]. Tailoring PA interventions at the *deep structure* level may enhance the acceptability and uptake of an intervention, which is expected to lead to greater improvements in PA when compared with a nontailored or surface-tailored PA intervention.

Objective

To address low PA levels, high cardiometabolic disease prevalence, and limitations of previous research using eHealth and mHealth technologies to promote PA among AA women, we developed *Smart Walk*, an 8-month, deep-structure culturally

tailored PA intervention for AA women aged 24 to 49 years (identifier: NCT02823379 [27]). Women in this age range were included in the intervention because it allowed us to tailor the PA intervention to the sociocultural norms and lifestyle factors of adult AA women of childbearing age and coincides with the age range of adults with the highest prevalence of smartphone use [28]. The purpose of this study was to address the relative lack of published reports documenting the development process of behavioral mHealth PA interventions by describing the development and initial usability testing results of *Smart Walk*.

Methods

Description of the Smart Walk App

The *Smart Walk* intervention is primarily delivered through the *Smart Walk* smartphone app. This app, developed specifically for the study and available for iOS and Android operating systems, includes 4 key features: (1) personal profile pages for participants to share personal information (ie, name, picture/image, age, city/neighborhood of residence, and brief biography), (2) theory-based multimedia PA promotion modules delivered weekly through brief videos and electronic text with images, (3) asynchronous discussion boards for participants to discuss each weekly PA module and give/receive social support, and (4) a PA self-monitoring tool that integrates with Fitbit Alta HR activity monitors for participants to monitor their daily and weekly PA levels. In addition to the information delivered through the smartphone app, participants received brief PA promotion text messages delivered 3 times per week. An overview of the key features of the *Smart Walk* intervention is provided in Table 1. Screenshots of the *Smart Walk* app feature are illustrated in Multimedia Appendices 1 and 2 and in a recently published paper [29] describing the rationale and design of the *Smart Walk* intervention.

Table 1. Overview of *Smart Walk* intervention components.

Intervention component	Description
Smartphone app features	
Personal profile pages	• Comparable with personal profile pages on commercially available social media websites (ie, Twitter and Facebook) • Allows participants to share information with other women in the study: name, personal picture/image, age, neighborhood/area of residence, and brief biographical narrative. • Aimed at creating a sense of community among study participants.
Weekly multimedia text and video modules	• Primary delivery channel for the educational and behavioral components of the intervention. • Modules consist of text- and image-based PA[a] promotion materials and brief 3- to 7-min videos describing the PA promotion topic of the week.
Discussion board forum with weekly discussion prompts	• Companion to weekly PA promotion modules • Provide a venue for participants to reflect on information presented, share their personal experiences about PA, and give/receive social support for PA. • Also includes a general Community Board forum, where participants share information and/or discuss topics that may not align with the weekly module topics. • Primary mechanism to foster social support for PA through dialog among participants
PA tracker	• Self-monitoring feature for tracking daily and weekly PA through interactive graphing functions • Integrates with the Fitbit Alta HR activity monitor
Text messages delivered 3 times per week	• Provide PA promotion reminders, tips, and encouragement

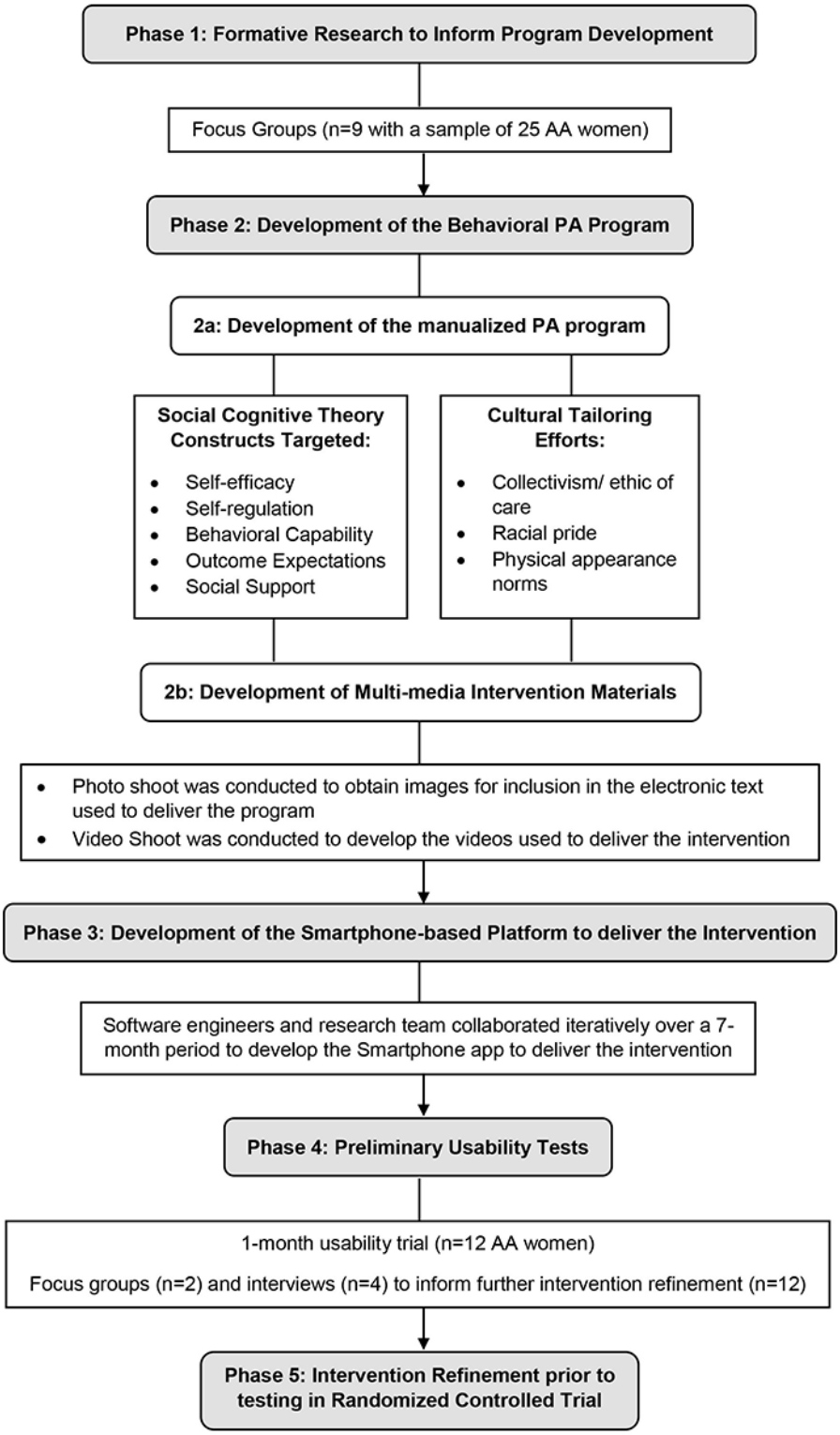
[a]PA: physical activity.

Results

Overview of the Smart Walk Development Process

The development of *Smart Walk* followed a 5-phase process. Phases 1 to 3 focused on intervention development, phase 4 involved usability testing, and phase 5 focused on intervention refinement based on usability testing results. Figure 1 provides an overview of these phases. Detailed descriptions of methods for each phase are presented below.

Figure 1. Overview of the Smart Walk development process.



Phase 1: Formative Research to Inform Theoretical and Cultural Relevance of the Smart Walk Intervention

In phase 1 of the intervention development, 25 AA women with obesity (mean age 38.5 years, SD 7.8; mean BMI 39.4 kg/m^2, SD 7.3) were recruited to participate in a series of 3 focus group sessions. These focus groups concentrated on collecting information on AA women's perceptions, experiences, and determinants related to PA to inform the cultural and theoretical relevance of the intervention. Eligibility criteria for participation included (1) self-identifying as an AA woman, (2) being between the ages of 24 and 49 years, (3) having a BMI greater than or equal to 30 kg/m^2, and (4) performing less than or equal to 60 min per week of moderate-to-vigorous intensity PA (according to the 2-item Exercise Vital Sign questionnaire [30]). These inclusion criteria represented the desired target population and end users of the *Smart Walk* intervention.

The study protocol, reported elsewhere [31], engaged each participant in a series of 3 focus group sessions. Each focus group session explored specific PA-related topics to inform development of the *Smart Walk* intervention. This cohort approach allowed us to collect in-depth information from participants while also being considerate of participant burden. The results of these focus group sessions described how 5 social cognitive theory constructs (ie, self-efficacy, self-regulation, behavioral capability, outcome expectations, and social support) could be integrated in the design of a culturally relevant PA intervention for AA women. They also served to identify how cultural norms and preferences of AA women could be incorporated into the intervention design and delivery. Previously published studies [31-33] provide an in-depth description of the methodology and results of these focus group sessions.

Phase 2: Development of the Behavioral Physical Activity Intervention and Wellness Contact Control Condition

Phase 2 of the intervention development was divided into 2 distinct subphases. The first subphase (phase 2a) focused on the development of the structured behavioral PA intervention as well as the attention-matched wellness control condition (named *Smart Health)* that would serve as a control group in the *Smart Walk* pilot trial. The second subphase (phase 2b) involved creation of the multimedia videos and print-based PA promotion materials used to deliver the PA intervention.

Phase 2a: Development of the Structured Physical Activity Promotion Program

In this phase, the research team used the qualitative data collected in phase 1 to refine and enhance a previously established PA promotion intervention for AA. This previously established program [34] was originally developed and pilot-tested in an 8-week randomized controlled trial using Facebook and text messages as the delivery method. Modifications to the PA program focused on (1) enhancing theoretical fidelity [20] of social cognitive theory constructs targeted by the intervention, (2) refining the cultural relevance, and (3) extending the length of the active intervention from 8 weeks to 4 months. The research team also identified and developed content for the wellness control during this phase. Health topics for the wellness control were selected based on their relevance to overall health and wellness but not inclusive of cardiometabolic disease risk (ie, breast and cervical cancer screenings, stress management, and preventing the cold and flu), including specific health topics relevant to AA women (ie, hair care and sickle cell disease). Content for the wellness contact control condition was developed through reviews of the scientific literature, rather than formative work with the study population.

Intervention content was developed collaboratively among members of the research team. These efforts resulted in the creation of 14 separate print-based PA promotion modules and 14 wellness modules to be delivered in the attention-matched control condition. Specific PA and health topics covered in these modules are presented in Table 2.

Table 2. Module topics by study group.

Module number	*Smart Walk* module topics	*Smart Health* module topics
1	Introduction to the national PA[a] guidelines and the health benefits of PA	Sunscreen and skin care
2	Overview of PA-related health disparities among African American women and the importance of being a PA role model	Hydration and water consumption
3	Time management and strategies for incorporating 30 min of PA into the day	Heat illness: causes, symptoms, and prevention
4	PA goal setting	Discussing your health with your doctor
5	Overcoming general barriers to PA	Breast cancer prevention and screening
6	Tips for increasing daily PA	Cervical cancer prevention and screening
7	Overcoming hair care barriers to PA	Hair care
8	Creating a social support network for PA	Oral health
9	Trying new types of activities	Eye health
10	Reducing sedentary time	Sickle cell disease
11	Dietary behaviors to complement PA	Stress management
12	Muscle strengthening and stretching activities to complement aerobic PA	Preventing the cold/flu
13	Dealing with setbacks	Staying healthy when traveling
14	Review of previous modules and maintenance of PA after the active intervention phase	Review of previous modules

[a]PA: physical activity.

After initial drafts of the print-based PA promotion and wellness attention-matched control group materials were developed, a professional copyeditor reviewed and edited the print-based intervention materials and developed oral scripts. Edits focused on grammar, message clarity, and simplifying the text to ensure all materials were written at or below an eighth-grade reading

level. Throughout the editing process, the copyeditor worked with the research team to ensure that the edits did not change the original meaning or cultural relevance of the intervention materials. Oral scripts included the same content as the print-based materials but were designed to be read out aloud by the study spokesperson during the development of the module videos used for the PA program and the wellness control condition.

Once finalized, the intervention was designed to target social cognitive theory constructs of behavioral capability, outcome expectations, social support, self-efficacy, and self-regulation and to address deep structure sociocultural characteristics of collectivism, racial pride/role modeling, and physical appearance norms of AA women (ie, hairstyle and body shape preferences of AA women). The strategies used to address these theoretical constructs and cultural characteristics are briefly described in Tables 3 and 4. Readers are also referred to a recent publication by our research team for a more in-depth description of these design characteristics [29].

Phase 2b: Development of the Multimedia Physical Activity and Health Promotion Modules

After the development of the structured PA promotion and attention-matched wellness control materials, the research team collaborated with a local photography/videography company to (1) create images to be included in the electronic print-based PA promotion materials and (2) develop the video vignettes used to deliver the intervention. To create the images for inclusion of the intervention materials, local AA women of diverse ages and body types were recruited to participate in a photo shoot. This photo shoot was designed to capture local AA women engaging in various types of PA. The decision to use local women in these intervention videos and images, as opposed to professional models or stock photos, was purposeful, as research [35,36] suggests that social modeling from real-life members of the community who possess characteristics of the study participants (eg, having a history of struggling with being physically active and trying to find a work and family life balance to allow being physically active) enhances the credibility and relatability of the information portrayed in a PA promotion program.

To create the 3- to 7-min intervention videos, 2 separate 1-day video shoots were conducted. At the first video shoot, initial versions of all module videos were filmed. After all videos were filmed, the videographer created rough cuts of the module videos for the research team to review. After an initial review of the videos, the research team determined several of the videos needed to be refilmed to enhance intervention delivery. After this reshoot, the videographer performed postproduction edits to all module videos and added background music, interactive graphics, and on-screen text. These postproduction edits served 2 functions: (1) to reinforce and emphasize key content discussed in the intervention and (2) to make the videos more interactive and engaging to watch. The research team worked iteratively with the videographer during this postproduction process until both parties were satisfied with final versions of all videos.

Table 3. Theoretical constructs targeted by the intervention.

Social cognitive theory construct	Brief description	How the construct is addressed in the intervention
Behavioral capability	Knowledge and skill to perform a PA[a]	Intervention materials provide information on the health benefits of regular PA, the national PA guidelines, how to gauge the intensity of PA performed, and strategies to achieve the national PA guidelines.
Outcome expectations	Anticipated outcomes of engaging in PA	Intervention materials highlight the health (ie, reduced cardiometabolic disease risk) and social benefits (ie, being a role model) of being physically active.
Social support	Extent to which significant referents approve, encourage, and/or influence performance of PA	Text messages provide messages of support and encouragement for PA. Discussion board forums prompt discussion focused on increasing social support for PA.
Self-regulation	Ability to manage social, cognitive, and motivational processes to achieve a desired PA goal	Participants are provided a static intervention goal of achieving a 150 min per week of PA and provided an activity monitor to self-monitor their PA to achieve this goal.
Self-efficacy	Confidence in oneself to take action and overcome barriers	Intervention messages provide encouragement for PA engagement (ie, verbal persuasion), intervention content includes images of AA[b] women engaging in various types of PA and PA testimonials (ie, social modeling), and participants are encouraged to track their PA as they implement behavior change strategies targeted by the intervention (ie, mastery experiences).

[a]PA: physical activity.

[b]AA: African American.

Table 4. Cultural characteristics targeted by the intervention.

Cultural characteristics	Brief description	How the cultural characteristic is addressed in the intervention
Collectivism	Prioritization of the needs of others (ie, family/close friends) before the needs of their own, which can result in AA[a] women reporting lack of time, energy, or resources for PA[b] engagement. Although this phenomenon is reported among women of other races/ethnicities, previous research, including our own formative work, has suggested this concept may be more accentuated in the AA community.	Intervention materials place emphasis on the following: • Importance of caretaking in the value system of AA women. • Regular PA engagement is an investment in the health and well-being of AA women, not taking *time away* from their family/friends or other responsibilities. • Regular PA will help AA women to perform their caretaking and other responsibilities with more energy and for a longer duration throughout the lifespan.
Racial pride/role modeling	Many AA women are aware of, and interested in, how their behaviors can contribute to the collective health and well-being of the AA community.	PA promotion materials emphasize that physically active AA women are positive role models for other members of the AA community, which can encourage others in their community (ie, family and friends) to become active and adopt healthy lifestyle behaviors.
Physical appearance preferences	Some AA women are hesitant to engage in PA because of the following reasons: • Perspiration can have a negative effect on their hairstyle. • They have the perception that PA will alter their desired body shape.	Intervention materials: • Include hairstyling strategies to reduce the negative effects of perspiration while performing PA (ie, use hair wraps and certain hair care products that negate the effects of sweating) and encourage women to adapt hairstyles that are less affected by perspiration (ie, braids and natural hairstyles). • Inform participants that engaging in PA at the levels recommended by the study (ie, 150 min per week) will not substantially change their body shape unless they also change their dietary habits. • Emphasize the health benefits of PA independent of weight loss (ie, reduced cardiometabolic disease risk, weight maintenance, and increased energy).

[a]AA: African American.

[b]PA: physical activity.

Phase 3: Development of the Smartphone-Based Platform to Deliver the Physical Activity Program

Phase 3 of program development focused on technical development of the smartphone app platform used to deliver the interventions. This task was accomplished through a collaborative process involving the study's principal investigator (PI) and a software engineer serving on the project team.

During initial meetings with the software development team, the PI provided a general overview of the operating system requirements (ie, available for iOS and Android devices), desired features, and preliminary thoughts on the visual layout of the study apps. Drawing on these initial discussions, the software engineer worked in 1- to 2-week iterative *sprints* over the course of a 7-month period to build the platform used to deliver both the PA and wellness attention-matched control interventions. During each sprint, the software engineer worked on a defined task until a prototype was ready for preliminary review and usability testing by the research team. Following review and testing by the research team, either further modifications were requested or the prototype was approved and filed as *complete*. This process was repeated over the course of 7 months until an initial version of both smartphone apps was developed.

The initial prototype of the *Smart Walk* intervention included the following 3 features: (1) multimedia PA promotion modules consisting of brief videos and electronic text with images, (2) discussion boards for participants to discuss the weekly PA modules and give/receive social support, and (3) a PA self-monitoring/tracking tool that integrated with the Fitbit Alta HR activity monitor. The *Smart Health* app included all the same features, with the exception of the PA self-monitoring/tracking feature. Screenshots of these initial prototypes are available upon request from the first author of the paper.

Phase 4: Preliminary Usability Test Trial

After the initial development of the study smartphone apps, a 1-month usability trial was conducted. The methods used in this trial allowed the research team to obtain participant feedback on (1) specific issues associated with the usability and functionality of the study apps and (2) their overall thoughts regarding health promotion programs when implemented in the real world. This trial also provided the opportunity for the research team to test the computer-based algorithms used to deliver the interventions (ie, uploading of new modules and discussion board topics and delivery of text messages) before large-scale testing.

Usability Trial Methods

A total of 12 insufficiently active (ie, ≤60 min per week of at least moderate PA according to the Exercise Vital Sign

questionnaire [30]) AA women with obesity (ie, BMI≥30 kg/m^2) who were between the ages of 24 and 49 years participated in this trial. These sample characteristics reflect the inclusion criteria for women subsequently recruited for a randomized pilot trial of the intervention. Participants were randomly assigned using stratified randomization based on participants' smartphone operating system (ie, iOS or Android) to receive either the culturally relevant PA program (ie, *Smart Walk)* or the wellness attention-matched control (*Smart Health*). This randomization method ensured balanced feedback from users of both operating systems on both study apps. After randomization, participants received all intervention materials originally designed to be delivered over a 4-month active intervention period in an abbreviated 1-month trial period.

During the trial, study staff communicated with participants (ie, via telephone, text message, or email based on participants' preferences) to inquire about any functionality or usability issues with the apps during weeks 1 and 3. Participants also occasionally emailed study staff and/or posted comments on the app discussion boards when they experienced a usability or functionality issue. After completion of the 1-month trial, participants were invited to participate in a focus group session to provide feedback on the smartphone-delivered program they received. Women not available to participate in the focus group session were provided the option to participate in a one-on-one in-person interview. Participants were provided US $30 for participating in the 1-month demonstration trial and an additional US $20 for participating in a focus group or interview session. No other strategies for recruitment or retention were employed.

Focus groups and interviews were led by an AA facilitator who was trained in qualitative data collection methods by the study PI. Guides used to facilitate focus group and interview sessions are presented in Textbox 1. All data collection sessions were audio-recorded, transcribed verbatim, and imported into NVivo12 (QSR International) for analysis. Content analysis [37] was used to analyze the participant narratives. In total, 2 coders reviewed study data independently and then met to discuss themes until consensus was reached. Final themes used in the analysis were based on major topics explored by the focus group/interview guide and repetitive themes that emerged during data collection assessments.

Textbox 1. Postusability trial focus group/interview guide questions. Similar guides were used in both study arms. The only differences between guides were (1) the reference to either Smart Walk or Smart Health, and (2) question 6 was only asked for participants assigned to the Smart Walk study group.

1. What are your overall thoughts about the [Smart Walk or Smart Health] app?
2. Please share your thoughts about the overall usability and functionality of the [Smart Walk or Smart Health] app.
3. Thinking about the multimedia video and text modules on the [Smart Walk or Smart Health] app, what are your thoughts on them?
4. What are your thoughts about the text messages you received?
5. Please describe your thoughts about the [Smart Walk or Smart Health] Discussion Boards?
6. Let's talk about the activity tracking feature. What are your thoughts on using it?
7. Overall, what are your opinions on the [Smart Walk or Smart Health] program as whole?
8. Other than a smartphone application, how would you like to receive a health promotion program?
9. Is there anything else you would like to tell us about the Smart Walk [or Smart Health] application or physical activity program that we have not already discussed?

Usability Trial Results

The 12 participants had a mean age of 35.0 years (SD 8.5) and a mean BMI of 40 kg/m^2 (SD 5.0). Of these 12 participants, 7 accessed the study smartphone apps with an Android device (n=3 in PA arm and n=4 in wellness arm), and 5 accessed the study smartphone apps with an iOS device (n=3 in PA arm and n=2 in wellness arm). Feedback regarding the usability and functionality of the *Smart Walk* app was obtained through 2 focus group sessions (group sizes were n=4 and n=2). Feedback on the *Smart Health* app was collected through 1 focus group (n=2) and 4 one-on-one interviews. The qualitative assessments lasted between 33 and 56 min for the focus group sessions and between 18 and 22 min for the one-on-one interviews. Participant narratives were classified into 3 overarching themes: usability/functionality issues, desire for enhanced personalization, and overall impressions of the intervention.

Usability/Functionality Issues

Participants identified several usability and functionality issues for the *Smart Walk* and *Smart Health* study apps. These concerns were classified into subthemes according to specific app feature. Participant quotes illustrating usability/functionality issues discussed below are presented in Multimedia Appendix 3, as are the refinements made to the intervention based on participant feedback.

Multimedia Modules

Overall, participants in both study groups reported encouraging sentiments regarding the information delivered through the multimedia modules. However, several usability and functionality issues associated with module delivery emerged. These included (1) excessive video buffering when participants had limited cellular service and/or intermittent wireless internet network connection; (2) audio portion of videos was not playing on select Android devices (this was because of an error in video filter settings); and (3) video display on iOS devices defaulting

to full screen, which did not allow participants to simultaneously listen to the video and read the text.

Discussion Board Forums

Participants identified 2 key usability/functionality issues associated with the app's discussion board feature. The first was a platform-specific issue for some, but not all, Android users that resulted in a deletion of unsaved text if participants rotated their phone while typing a discussion board post. The second issue was related to participant desire for enhanced interactivity. Participants indicated that they would like to receive real-time notifications when other study participants posted on the discussion boards, as opposed to having to periodically check the study app to see if anyone posted and/or replied to the boards. When discussing strategies to provide these notifications, text messages and push notifications were mentioned by participants. As the conversations progressed, it became clear that most participants preferred the use of push notifications over text messages.

Activity Tracker

Participant feedback on the activity tracker feature revealed 2 primary concerns. The first issue was related to the wrist-worn Fitbit activity monitor not accurately recording all moderate-to-vigorous activities. Several participants noted engaging in physical activities with restricted arm movement (ie, stationary cycling) that were not accurately being recorded by the wrist-worn device. Similarly, a few participants perceived this to be an issue, when in actuality, they were engaging in activities not considered moderate-to-vigorous intensity aerobic activities (ie, yoga and walking at a cadence that did not meet the criterion for moderate-intensity activity). The second issue was that, for some participants, the commercial Fitbit software was not automatically communicating with the *Smart Walk* app. This technological issue resulted in the *Smart Walk* app not always updating participants' minutes per day of moderate-to-vigorous PA unless they first opened and refreshed the commercial Fitbit app.

Desire for Enhanced Personalization and Individual-Level Tailoring

Throughout the focus groups and individual interviews, participants expressed the desire for both study smartphone apps to include additional features to promote an enhanced sense of personalization and individual-level tailoring. Participants described that although they viewed the intervention content as favorable, additional steps could be taken to facilitate a sense of community among study participants and to individually tailor the program to each participant. Discussions on this topic resulted in participants suggesting 3 key modifications to the study smartphone apps: (1) addition of personal profile pages, (2) personalization of the discussion boards, and (3) individual-level tailoring of text messages. Participant quotes highlighting these suggested revisions to the study app are presented in Mutlimedia Appendix 3.

Addition of Personal Profile Pages

One strategy that seemed to enhance personalization of the apps was the addition of a personal profile feature. This feature was initially brought up in the first *Smart Walk* focus group session. Subsequent conversations regarding the creation of personal profile pages revealed that participants envisioned this feature to be similar to profile pages on social media websites (ie, Facebook and Twitter) and expressed that adding this feature would help create a stronger sense of community among users. Specific information recommended by participants to include on this feature included participant name, personal picture/image, city/neighborhood of residence, and a brief biography. However, it should be noted that not all participants were as enthusiastic about the idea of adding a personal profile page to the study apps. In the end, there was a consensus among participants that the profile page would be a good addition if it was optional for participants to complete.

Personalized Discussion Board Forums

Another topic that emerged was participants' desire for additional discussion board forums where they could create and/or drive the conversation narrative, rather than relying on the weekly topic-specific discussion board prompts designed to facilitate discussion. Discussion board forums during the usability trial were tied directly to the weekly module topics and included prompts to facilitate topic-specific discussion. Some participants alluded to the notion that this felt restrictive and did not provide an opportunity to engage in organic dialog on topics not directly related to the weekly module topics. Participants indicated that they would like the opportunity to create their own discussion threads and discuss topics not specifically related to the weekly modules.

Individual-Level Tailoring of Text Messages

Discussion on the study text messages revealed 2 minor modifications that could be implemented to help achieve a sense of individual-level tailoring: (1) address participants by name in the text messages and (2) allow participants to specify the time of day they would like to receive text messages. Several participants noted that they perceived the text messages as impersonal or generic. These women emphasized that including their name in the text messages would make them feel *special* and provide a sense of individual-level tailoring. With reference to the time of day the text messages were delivered, the study protocol during the demonstration trial had all participants receive a text message at 8:30 am. Several participants stated that this time was not ideal, and they would prefer to select the time of day to receive the text messages.

Overall Impressions of the Intervention

Participants reported favorable overall impressions of the smartphone-delivered interventions. Quotes highlighting this sentiment for participants assigned to the PA intervention (ie, *Smart Walk*) included, "The app itself, I thought it was really informative, and it was fun to watch the videos, and participate in the discussion boards," "I learned new stuff," and "I thought that it was cool... And it was clear. So I enjoyed it." Similarly, participants assigned to the overall health and wellness group (ie, *Smart Health*) enjoyed the app and felt it was useful for others. Despite favorable overall impressions of the intervention, several participants noted the abbreviated 1-month duration of the trial resulted in too much information being delivered over a short period (ie, 1-2 text messages each day and new module

and discussion board prompt every 2-3 days). Quotes highlighting this included "I think the text messages were a bit excessive." and "Since it was condensed, we go text more often than we normally would...I had to remember that. I'm like, this is a little, I can't keep up." As conversations on this topic continued and the facilitator described the delivery schedule for forthcoming randomized controlled trial (ie, 1 new module topic/discussion board prompt per week and 3 text messages each week during the active 4-month intervention phase), participants expressed that the intervention dose for which materials were originally designed to be delivered seemed more appropriate.

Phase 5: Intervention Refinement

Phase 5 focused on refining the intervention according to participant feedback from phase 4. To accomplish this, the research team systematically evaluated participant narratives and recommendations for intervention improvement. On the basis of this analysis, numerous refinements and modifications were made to the study apps, which are illustrated in Multimedia Appendix 3. The screenshots of the final versions of both the *Smart Walk* and *Smart Health* apps are presented in Multimedia Appendices 1 and 2.

Discussion

Principal Findings

The development of the *Smart Walk* interventions occurred over a 3-year period. It was made possible through a National Institutes of Health (NIH) career development award (ie, K99HL129012 and R00HL129012) and involved expertise from a diverse set of researchers, software engineers, industry professionals (ie, videographer/photographer and copyeditor), and intended users of the intervention. The methods used emphasize the importance of creating a transdisciplinary research team when developing innovative approaches to address major public health issues and the significance of conducting extensive formative research with end users of an intervention.

A novel aspect of the information reported in this paper is the methods and results of the 1-month usability trial (ie, phase 4 of intervention development). The methods employed in this phase allowed the research team to obtain feedback on the usability and functionality of the study apps as well as participants' overall impressions of the interventions when implemented in a real-world setting. We selected this method of participant usability testing, as opposed to more preliminary laboratory-based methods (ie, user-centered design or *think aloud* techniques), because the research team conducted these types of laboratory tests internally during phase 3 of intervention development. Given the research team included commercial mHealth app users who also have experience developing research-based eHealth and mHealth PA interventions [34,38,39], we relied on our previous experience and expertise in these areas to conduct preliminary usability tests.

The results of the usability trial emphasize the importance of pilot-testing smartphone apps on various operating systems (ie, iOS and Android) and smartphone devices, as our results indicated app functionality varied based on these factors (ie, video sound was not audible on select Android devices and defaulted to full screen on iOS devices). Another important aspect that emerged from the usability trial was the need to enhance the individual-level tailoring and/or personalization of the interventions. Participants discussed, at length, the need for the interventions to be tailored to them at the individual level and not just to their sociocultural characteristics associated with being an AA woman. The strategies proposed by participants to achieve this level of personalization included incorporating their names into the app features (ie, home page and activity tracker) and text messages, allowing them to create personal profiles, and creating a community discussion board forum where they can drive the discussion narrative, rather than the research team. Participants also noted the need to enhance the interactivity of the app discussion boards by adding push notifications to notify participants when other women post on the discussion boards. Participants described that this feature is common among other apps they use (ie, Facebook) and is expected if a goal of the study was to promote communication and dialog with other users.

The results of the usability trial also highlight the challenges researchers face when developing mHealth interventions. As evidenced by our usability trial results, participants expect research-based smartphone apps to be just as interactive, engaging, and easy to use as commercially available apps. In our experience, this can be somewhat difficult to achieve because of grant funding cycles (ie, typically no more than 2-5 years to develop, implement, and evaluation outcomes), available resources (ie, money and access to software engineers), and the rapid pace at which technology changes. In addition, there is a need for researchers to justify why creation of a new app is necessary, as there are thousands of other health promotion apps commercially available to consumers. In our case, the novelty of the intervention is that it is driven by theory and deep structure and culturally tailored to the sociocultural norms and behavior preferences of AA women, which is an important addition of the eHealth and mHealth field of study. These factors are designed to enhance acceptability and salience of the intervention, which, in turn, is expected to lead to the desired behavioral outcome of increased PA.

It is also important to discuss the nuanced relationship between the behavioral PA program and the smartphone app used to deliver it, as we view these as complementary yet separate entities. The behavioral PA program, developed in phases 1 and 2, can be adapted for delivery using other modalities (eg, in-person group-based sessions) without significantly altering cultural relevance or theoretical underpinnings of the program. Similarly, the smartphone app used to deliver the PA program can be modified to deliver other health promotion programs (ie, as was done for the *Smart Health* attention-matched wellness control condition). Given the differentiation between the behavioral PA program and the delivery channel can sometimes be a source of confusion when discussing eHealth and mHealth research with colleagues, community members, and students alike, we encourage researchers conducting future eHealth and mHealth interventions to clearly distinguish between the behavioral components of a program and the delivery channel when applicable.

Future Directions

Smart Walk is currently being tested in an 8-month randomized controlled pilot trial. Cardiometabolic end points of the intervention include reductions in total cholesterol, low-density lipoprotein cholesterol, systolic and diastolic blood pressure, fasting glucose, aortic pulse wave velocity, and proinflammation cytokines (tumor necrosis factor alpha and interleukin 1 beta); an increase in high-density lipoprotein cholesterol; and improvements in cardiorespiratory fitness.

Data collected from this trial will provide information on preliminary efficacy of the intervention and highlight areas for further intervention refinement before large-scale testing.

Conclusions

Smart Walk is an innovative approach to promote PA and reduce cardiometabolic disease risk among AA women. The formative work described in this paper was designed to enhance the cultural and theoretical relevance of the intervention as well as optimize user experience with using the Smart Walk app. The methods described can provide a framework for other researchers to follow when developing future eHealth and mHealth interventions.

Acknowledgments

The development of *Smart Walk* was made possible through postdoctoral training and early career development grants awarded to the PI by the NIH/National Heart, Lung, and Blood Institute (K99HL129012 and R00HL129012). The research team is thankful to the funding agencies, community members, and participants who have contributed and supported the development of this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of the Smart Walk smartphone app [28].
[PDF File (Adobe PDF File), 1016 KB - mhealth_v8i3e15346_app1.pdf]

Multimedia Appendix 2
Screenshots of the Smart Health smartphone app [28].
[PDF File (Adobe PDF File), 438 KB - mhealth_v8i3e15346_app2.pdf]

Multimedia Appendix 3
Modifications and refinements to the intervention based on usability trial results.
[DOCX File , 17 KB - mhealth_v8i3e15346_app3.docx]

References

1. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, American Heart Association Statistics CommitteeStroke Statistics Subcommittee. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. Circulation 2017 Mar 07;135(10):e146-e603 [FREE Full text] [doi: 10.1161/CIR.0000000000000485] [Medline: 28122885]
2. Gaesser GA, Angadi SS, Sawyer BJ. Exercise and diet, independent of weight loss, improve cardiometabolic risk profile in overweight and obese individuals. Phys Sportsmed 2011 May;39(2):87-97. [doi: 10.3810/psm.2011.05.1898] [Medline: 21673488]
3. Colberg SR, Sigal RJ, Fernhall B, Regensteiner JG, Blissmer BJ, Rubin RR, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement executive summary. Diabetes Care 2010 Dec;33(12):2692-2696 [FREE Full text] [doi: 10.2337/dc10-1548] [Medline: 21115771]
4. 2008 physical activity guidelines for Americans. Hyattsville, MD: US Department of Health and Human Services; 2008.
5. Williams WM, Yore MM, Whitt-Glover MC. Estimating physical activity trends among blacks in the United States through examination of four national surveys. AIMS Public Health 2018;5(2):144-157 [FREE Full text] [doi: 10.3934/publichealth.2018.2.144] [Medline: 30094277]
6. Ogden CL, Carroll MD, Flegal KM. Prevalence of obesity in the United States. J Am Med Assoc 2014 Jul;312(2):189-190. [doi: 10.1001/jama.2014.6228] [Medline: 25005661]
7. Writing Group Members, Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, American Heart Association Statistics Committee, Stroke Statistics Subcommittee. Heart disease and stroke statistics-2016 update: a report from the American Heart Association. Circulation 2016 Jan 26;133(4):e38-360. [doi: 10.1161/CIR.0000000000000350] [Medline: 26673558]

8. Centers for Disease ControlPrevention (CDC). 2014. Summary Health Statistics: National Health Interview Survey URL: http://www.cdc.gov/nchs/nhis/shs/tables.htm [accessed 2019-01-16]
9. Carnethon MR, Pu J, Howard G, Albert MA, Anderson CA, Bertoni AG, American Heart Association Council on EpidemiologyPrevention; Council on Cardiovascular Disease in the Young; Council on CardiovascularStroke Nursing; Council on Clinical Cardiology; Council on Functional GenomicsTranslational Biology;Stroke Council. Cardiovascular health in African Americans: a scientific statement from the American Heart Association. Circulation 2017 Nov 21;136(21):e393-e423. [doi: 10.1161/CIR.0000000000000534] [Medline: 29061565]
10. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
11. Müller AM, Alley S, Schoeppe S, Vandelanotte C. The effectiveness of e-& mHealth interventions to promote physical activity and healthy diets in developing countries: A systematic review. Int J Behav Nutr Phys Act 2016 Oct 10;13(1):109 [FREE Full text] [doi: 10.1186/s12966-016-0434-2] [Medline: 27724911]
12. 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]
13. Buchholz SW, Wilbur J, Ingram D, Fogg L. Physical activity text messaging interventions in adults: a systematic review. Worldviews Evid Based Nurs 2013 Aug;10(3):163-173. [doi: 10.1111/wvn.12002] [Medline: 23746267]
14. Beishuizen CR, Stephan BC, van Gool WA, Brayne C, Peters RJ, Andrieu S, et al. Web-based interventions targeting cardiovascular risk factors in middle-aged and older people: a systematic review and meta-analysis. J Med Internet Res 2016 Mar 11;18(3):e55 [FREE Full text] [doi: 10.2196/jmir.5218] [Medline: 26968879]
15. Hakala S, Rintala A, Immonen J, Karvanen J, Heinonen A, Sjögren T. Effectiveness of physical activity promoting technology-based distance interventions compared to usual care. Systematic review, meta-analysis and meta-regression. Eur J Phys Rehabil Med 2017 Dec;53(6):953-967. [doi: 10.23736/S1973-9087.17.04585-3] [Medline: 28466628]
16. Feter N, Dos Santos TS, Caputo EL, da Silva MC. What is the role of smartphones on physical activity promotion? A systematic review and meta-analysis. Int J Public Health 2019 Jun;64(5):679-690. [doi: 10.1007/s00038-019-01210-7] [Medline: 30758514]
17. Martinez SM, Arredondo EM, Perez G, Baquero B. Individual, social, and environmental barriers to and facilitators of physical activity among Latinas living in San Diego County: focus group results. Fam Community Health 2009;32(1):22-33. [doi: 10.1097/01.FCH.0000342814.42025.6d] [Medline: 19092432]
18. Attwood S, Morton KL, Mitchell J, Van Emmenis M, Sutton S, Programme Team VBI. Reasons for non-participation in a primary care-based physical activity trial: a qualitative study. BMJ Open 2016 May 23;6(5):e011577 [FREE Full text] [doi: 10.1136/bmjopen-2016-011577] [Medline: 27217288]
19. Ingram D, Wilbur J, McDevitt J, Buchholz S. Women's walking program for African American women: expectations and recommendations from participants as experts. Women Health 2011 Aug 31;51(6):566-582 [FREE Full text] [doi: 10.1080/03630242.2011.606357] [Medline: 21973111]
20. Rovniak LS, Hovell MF, Wojcik JR, Winett RA, Martinez-Donate AP. Enhancing theoretical fidelity: an e-mail-based walking program demonstration. Am J Health Promot 2005;20(2):85-95. [doi: 10.4278/0890-1171-20.2.85] [Medline: 16295700]
21. Centers for Disease Control and Prevention, Division of Nutrition, Physical Activity, and Obesity, National Center for Chronic Disease Prevention and Health Promotion. 2019. Nutrition, Physical Activity, and Obesity: Data, Trends and Maps URL: https://www.cdc.gov/nccdphp/dnpao/data-trends-maps/index.html
22. Blackwell DL, JW. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. 2015. Tables of Summary Health Statistics for US AdultsNational Health Interview Survey URL: https://ftp.cdc.gov/pub/Health_Statistics/NCHS/NHIS/SHS/2014_SHS_Table_A-14.pdf [accessed 2017-08-31]
23. Pew Research Center. 2017. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2017-08-31]
24. Pew Research Center. 2018. Social Media Use in 2018 URL: https://www.pewinternet.org/2018/03/01/social-media-use-in-2018/ [accessed 2019-06-28]
25. Joseph RP, Royse KE, Benitez TJ. A systematic review of electronic and mobile health (e- and mHealth) physical activity interventions for African American and Hispanic women. J Phys Act Health 2019 Mar 01;16(3):230-239. [doi: 10.1123/jpah.2018-0103] [Medline: 30782040]
26. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. Ethn Dis 1999;9(1):10-21. [Medline: 10355471]
27. US National Library of Medicine. Clinicaltrials.gov URL: https://www.clinicaltrials.gov/ct2/home
28. Pew Research Center. 2019. Mobile Fact Sheet URL: https://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-06-28]
29. Joseph RP, Ainsworth BE, Vega-López S, Adams MA, Hollingshead K, Hooker SP, et al. Rationale and design of Smart Walk: a randomized controlled pilot trial of a smartphone-delivered physical activity and cardiometabolic risk reduction intervention for African American women. Contemp Clin Trials 2019 Feb;77:46-60. [doi: 10.1016/j.cct.2018.12.011] [Medline: 30576840]

30. Coleman KJ, Ngor E, Reynolds K, Quinn VP, Koebnick C, Young DR, et al. Initial validation of an exercise "vital sign" in electronic medical records. Med Sci Sports Exerc 2012 Nov;44(11):2071-2076. [doi: 10.1249/MSS.0b013e3182630ec1] [Medline: 22688832]
31. Joseph RP, Ainsworth BE, Mathis L, Hooker SP, Keller C. Utility of social cognitive theory in intervention design for promoting physical activity among African-American women: a qualitative study. Am J Health Behav 2017 Sep 01;41(5):518-533 [FREE Full text] [doi: 10.5993/AJHB.41.5.1] [Medline: 28760174]
32. Joseph RP, Ainsworth BE, Mathis L, Hooker SP, Keller C. Incorporating religion and spirituality into the design of community-based physical activity programs for African American women: a qualitative inquiry. BMC Res Notes 2017 Oct 23;10(1):506 [FREE Full text] [doi: 10.1186/s13104-017-2830-3] [Medline: 29058603]
33. Joseph RP, Coe K, Ainsworth BE, Hooker SP, Mathis L, Keller C. Hair as a barrier to physical activity among African American women: a qualitative exploration. Front Public Health 2017;5:367 [FREE Full text] [doi: 10.3389/fpubh.2017.00367] [Medline: 29404317]
34. Joseph RP, Keller C, Adams MA, Ainsworth BE. Print versus a culturally-relevant Facebook and text message delivered intervention to promote physical activity in African American women: a randomized pilot trial. BMC Womens Health 2015 Mar 27;15:30 [FREE Full text] [doi: 10.1186/s12905-015-0186-1] [Medline: 25886945]
35. Durant NH, Joseph RP, Cherrington A, Cuffee Y, Knight B, Lewis D, et al. Recommendations for a culturally relevant internet-based tool to promote physical activity among overweight young African American women, Alabama, 2010-2011. Prev Chronic Dis 2014 Jan 16;11:130169 [FREE Full text] [doi: 10.5888/pcd11.130169] [Medline: 24433625]
36. Pekmezi D, Marcus B, Meneses K, Baskin ML, Ard JD, Martin MY, et al. Developing an intervention to address physical activity barriers for African-American women in the deep south (USA). Womens Health (Lond) 2013 May;9(3):301-312 [FREE Full text] [doi: 10.2217/whe.13.20] [Medline: 23638785]
37. 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]
38. Joseph RP, Pekmezi D, Dutton GR, Cherrington AL, Kim Y, Allison JJ, et al. Results of a culturally adapted internet-enhanced physical activity pilot intervention for overweight and obese young adult African American women. J Transcult Nurs 2016 Mar;27(2):136-146 [FREE Full text] [doi: 10.1177/1043659614539176] [Medline: 24934566]
39. Joseph RP, Pekmezi DW, Lewis T, Dutton G, Turner LW, Durant NH. Physical activity and Social Cognitive Theory outcomes of an internet-enhanced physical activity intervention for African American female college students. J Health Dispar Res Pract 2013;6(2):1-18 [FREE Full text] [Medline: 25215265]

Abbreviations

AA: African American
eHealth: electronic health
mHealth: mobile health
NIH: National Institutes of Health
PA: physical activity
PI: principal investigator

Edited by G Eysenbach; submitted 03.07.19; peer-reviewed by G Signorelli, A Kinsey; comments to author 03.10.19; revised version received 07.11.19; accepted 16.12.19; published 02.03.20.

Please cite as:
Joseph RP, Keller C, Vega-López S, Adams MA, English R, Hollingshead K, Hooker SP, Todd M, Gaesser GA, Ainsworth BE
A Culturally Relevant Smartphone-Delivered Physical Activity Intervention for African American Women: Development and Initial Usability Tests of Smart Walk
JMIR Mhealth Uhealth 2020;8(3):e15346
URL: https://mhealth.jmir.org/2020/3/e15346
doi:10.2196/15346
PMID:32130198

Original Paper

Quality of Life of Women After a First Diagnosis of Breast Cancer Using a Self-Management Support mHealth App in Taiwan: Randomized Controlled Trial

I-Ching Hou[1*], RN, PhD; Hsin-Yi Lin[1,2*], RN, MS; Shan-Hsiang Shen[3], PhD; King-Jen Chang[4], MD, PhD; Hao-Chih Tai[5], MD, PhD; Ay-Jen Tsai[4], BA; Patricia C Dykes[6,7], RN, MA, PhD

[1]School of Nursing, National Yang-Ming University, Taipei City, Taiwan

[2]Department of Nursing, Hsinchu Mackay Memorial Hospital, Hsinchu, Taiwan

[3]Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan

[4]Taiwan Breast Cancer Foundation, Taipei, Taiwan

[5]Department of Surgery, National Taiwan University, Taipei, Taiwan

[6]Center for Patient Safety Research and Practice, Brigham and Women's Hospital, Boston, MA, United States

[7]Harvard Medical School, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
I-Ching Hou, RN, PhD
School of Nursing
National Yang-Ming University
Number 155, Section 2, Linong Street, Beitou District
Taipei City, 11221
Taiwan
Phone: 886 28267000 ext 7315
Email: evita@ym.edu.tw

Abstract

Background: There are over 2 million newly diagnosed patients with breast cancer worldwide with more than 10,000 cases in Taiwan each year. During 2017-2018, the National Yang-Ming University, the Taiwan University of Science and Technology, and the Taiwan Breast Cancer Prevention Foundation collaborated to develop a breast cancer self-management support (BCSMS) mHealth app for Taiwanese women with breast cancer.

Objective: The aim of this study was to investigate the quality of life (QoL) of women with breast cancer in Taiwan after using the BCSMS app.

Methods: After receiving a first diagnosis of breast cancer, women with stage 0 to III breast cancer, who were recruited from social networking sites or referred by their oncologists or oncology case managers, were randomized 1:1 into intervention and control groups. Intervention group subjects used the BCSMS app and the control group subjects received usual care. Two questionnaires—the European Organization for Research and Treatment of Cancer (EORTC) Quality-of-Life Questionnaire Core 30 (QLQ-C30) and the EORTC Breast Cancer-Specific Quality-of-Life Questionnaire (QLQ-BR23)—were distributed to subjects in both arms. Paper-based questionnaires were used at baseline; paper-based or Web-based questionnaires were used at 1.5-month and 3-month follow-up evaluations. All evaluations were self-assessed and anonymous, and participants were blinded to their allocation groups. Descriptive analysis, the Pearson chi-square test, analysis of variance, and the generalized estimating equation were used to analyze the data. Missing values, with and without multi-imputation techniques, were used for sensitivity analysis.

Results: A total of 112 women were enrolled and randomly allocated to either the experimental group (n=53) or control group (n=59). The follow-up completion rate was 89.3% (100/112). The demographic data showed homogeneity between the two groups in age (range 50-64 years), breast cancer stage (stage II), marital status (married), working status (employed), and treatment status (receiving treatments). The mean total QoL summary scores from the QLQ-C30 (83.45 vs 82.23, *P*=.03) and the QLQ-BR23 (65.53 vs 63.13, *P*=.04) were significantly higher among the experimental group versus the control group, respectively, at 3 months.

Conclusions: This research provides support for using a mobile health care app to promote the QoL among women in Taiwan after a first diagnosis of breast cancer. The BCSMS app could be used to support disease self-management, and further evaluation of whether QoL is sustained is warranted.

Trial Registration: ClinicalTrials.gov NCT004174248; https://clinicaltrials.gov/ct2/show/NCT04174248

(JMIR Mhealth Uhealth 2020;8(3):e17084) doi:10.2196/17084

KEYWORDS

breast cancer; mHealth app; self-management; quality of life

Introduction

Background

Breast cancer has been the most common type of malignant cancer in women in Taiwan for more than 10 years and is the fourth-leading cause of death [1]. Each year, more than 10,000 women are diagnosed with breast cancer, and about 2000 women die from breast cancer: in 2014, there were 12,714 new diagnoses and 2083 deaths. The incidence of breast cancer in Taiwan is common in younger women, 40-65 years of age, with stages 0 to II being the most common at diagnosis [2]. With early detection and treatment of breast cancer, the treatment effect is good. With appropriate treatment, the 5-year relative survival rate approaches 87.2% and breast cancer is often regarded as a chronic disease. Breast cancer treatment includes breast preservation or resection, chemotherapy, hormone therapy, and targeted therapy [3-5]. The most common side effects differ by treatment method and include surgical-wound pain, nausea, and vomiting (see Table 1). Treatment side effects can cause anxiety and depression, which can affect a woman's willingness to undergo treatment and can severely decrease the quality of life (QoL) for women with breast cancer [6-12].

Today, the *self-management* model is generally valued and promoted in Taiwan and abroad [13-15]. Self-management is different from traditional disease management; this includes education, emphasizing patient-centered and disease-oriented self-management, and taking the initiative to participate in health care activities. With self-management, the patient learns problem solving, disease control, life adjustment, physical and mental symptom management, and lifestyle changes in order to coexist with a chronic disease in daily life. A systematic review on chronic disease intervention studies showed that self-management can significantly improve knowledge of chronic disease, enhance self-care behavior, and increase self-efficacy [16]. In a study of chronic disease self-management applied to women with breast cancer, it was found that self-management measures can effectively improve the QoL for women living with breast cancer, increase self-efficacy, and help women to better manage their medical and emotional tasks [17].

The World Health Organization defines QoL as "an individual's perception of their position in life in the context of the culture and value systems in which they live in relation to their goals, expectations, standards and concerns. It is a multifaceted concept affected by the person's physical health, psychological state, personal beliefs, social relationships and their relationship to salient features of their environment" [18]. QoL can be objectively measured. While breast cancer is the most common cancer among Taiwanese women, the survival rate is high. Providing patients with adequate physical and psychological treatment and care could improve their QoL, which becomes an important indicator of the quality of medical care [19-21]. Therefore, the measurement of QoL can objectively reflect the response of women with breast cancer to disease and provide important evaluation indicators for the effectiveness of interventions.

The use of mobile apps by patients with cancer is becoming more common. More than half of patients with cancer are willing to transmit information through an app to support their treatment, and 84% of medical professionals, mainly physicians, support the use of apps among this population [22,23]. One study suggests that apps can support outpatient visits, documentation of adverse events, treatment, and medication reminders [22]. Mobile apps have also been developed for breast cancer self-management [24-27]. However, we are unaware of any mobile self-management apps developed specifically for Taiwanese women with breast cancer. Therefore, our team developed the breast cancer self-management support (BCSMS) mHealth app to address this gap. The development and usability testing of the BCSMS app has been reported previously (publication is forthcoming). The purpose of this study was to evaluate the impact of using the BCSMS app on the QoL of Taiwanese women after an initial diagnosis of breast cancer.

Table 1. Common side effects of breast cancer treatments, which differ by treatment method.

Treatment	Common side effects
Surgery	Surgical-wound pain and lymphedema
Chemotherapy	Nausea, vomiting, loss of appetite, oral mucositis, fatigue, hair loss, myelosuppression, neuropathy cognitive disorders, weight gain, and poor memory
Radiotherapy	Dermatitis, sleep disturbance, and fatigue
Hormone therapy	Menopause, lack of libido, hot flashes, headache, night sweats, and insomnia

Prior Work

In Taiwan, since June 2017, the National Yang-Ming University, the Taiwan University of Science and Technology, and the Taiwan Breast Cancer Prevention and Research Foundation collaborated to develop the BCSMS app. Eight main features of the app included the following: (1) the evidence or *knowledge* about breast cancer, (2) *exercise and rehabilitation* after surgery, (3) *diet and nutrition* for breast cancer patients, (4) *emotional support* to prevent anxiety and depression, (5) a *personal health record* for tracking treatment and side effects, (6) *social resource* information, (7) *experience sharing*, and (8) *expert consulting*. We pilot-tested the BCSMS app with 45 Taiwanese women with breast cancer in 2018 using the modified *technology acceptance model of mobile services* survey. From this pilot test, our team found that the BCSMS app had sound usability and was accepted by the participants. The BCSMS app has since been outsourced to a technology company for long-term maintenance and can be downloaded for widespread use. To access the app, the keyword “ibreast” should be used to search for the iOS version and the keywords “pink passport” should be used for the Android version.

Methods

Study Design and Setting

The study was a single-blinded, parallel-group, randomized controlled trial with a pretest evaluation (T0), as well as 1.5-month (T1) and 3-month (T2) follow-up evaluations. The study sites included two medical centers— National Taiwan University Hospital and MacKay Memorial Hospital—and one area hospital—Hsinchu Mackay Memorial Hospital—in northern Taiwan. Every patient at each of the study sites received similar care from their health care team (eg, oncology clinic follow-up every 3 weeks during chemotherapy; oncology nurse case managers available to work face-to-face or via telephone with patients when needed on their cancer journey; and transfers by their physicians for consultations with health care professionals in various disciplines, such as pharmacists, physical therapists, nutritionists, psychologists, and social workers, when they had related health problems). The health care professionals’ care models influencing QoL improvement were assumed to be similar in this study.

Ethical Considerations

This study followed the ethical principles of the Declaration of Helsinki [28] and was approved by the Institutional Review Board (IRB) of National Yang Ming University in Taiwan (IRB No. YM107109E). Participation in the study was voluntary. Eligible participants were provided with two different informed consent forms according to their assigned groups. The consent forms included the same information about the protocol regarding their QoL data collection, and every participant was provided US $3 for each pretest evaluation and follow-up evaluation after the study. The information items that differed on the experimental group’s consent form were (1) a brief description about the contents of the BCSMS app and (2) a statement communicating that the BCSMS app would be installed on participants’ mobile phones with remote technology support. The protocol was determined to be of minimal risk to the participants. Nevertheless, all participant data were anonymized and stored on an encrypted, password-protected server.

Recruiting the Study Participants, Sample Size, and Randomization

Women with breast cancer who met the following inclusion criteria were recruited: (1) first diagnosis of stage 0 to III breast cancer within the past year, (2) aged 20-65 years, to avoid barriers with respect to aging influencing mobile health usability for older adults [29], (3) had an Android or iOS mobile phone, (4) able to read and write in Chinese, and (5) willing to participate in the study and provide informed consent.

A priori power analysis was calculated using G*Power, version 3.1 (Heinrich-Heine-Universität Düsseldorf), by performing F tests and repeated-measures, between-subjects factor analysis of variance; a medium effect size of 0.25, a significance level of .05, and a power of 0.8 were used for sample size calculation according to Cohen [30]. The a priori sample size was 41 in each group. We oversampled our eligible patient population to account for potential dropouts and expected to include a total of 106 patients in this study.

Participants were recruited in two different ways. First, a recruitment ad was posted on the Taiwan Cancer Foundation social networking sites (eg, Facebook and Line) and interested participants contacted us through the online registration. Second, patients were referred by their oncologists and oncology case managers from the study settings. To blind the study group participants to allocation [31] and prevent selection bias (eg, technology novelty bias) [32], the ad stated that participant recruitment was for a QoL evaluation following cancer treatment only, with no mention of the BCSMS app. The recruitment poster was used by recruiters to introduce the patients to the purpose of the study. For homogeneity of patients at each stage of breast cancer—stage 0 to III—between the two groups, the expected numbers of patients with stage 0, stage I, stage II, and stage III breast cancer were about 20, 36, 36, and 14, respectively: the incidence rate of each stage [1] was used and multiplied by 106. Next, participants at each stage were randomly assigned 1:1 into one of two study groups: control or experimental. The randomization scheme was generated using the website Randomization.com [33] for participants at each stage.

Data Collection Procedures

The study team collected data from January to July 2019. Patients were blinded to their allocation groups—experimental or control—and their pretest data were collected via a paper-based instrument. After the pretest data were collected, the BCSMS app was installed on the mobile phones of the participants in the experimental group and they were taught how to use the app. Each participant in the experimental group could use BCSMS app any time as needed; there were no prompts or reminders from the study team. The entire pretest study design was done via face-to-face contact at the study sites. Data collection for the two follow-up evaluations was completed via a Web-based instrument, together with a phone call, email, and/or communication software (eg, Line). Paper-based

instruments were also provided for participants that chose not to use the Web-based instrument. All evaluations were anonymous. During data collection, the content of the BCSMS app was frozen, however, the intervention group was supported in overcoming any technical problems associated with the BCSMS app (eg, log-in problems and data entry problems). During the study, participants in both groups received the same care, concurrently, from health care professionals at the study sites.

Instruments

The QoL instrument consisted of two parts. The first part included demographic items (eg, age, disease stage, marital status, and working status) and treatment-related items. The second part was the Taiwan Chinese version of two Quality-of-Life Questionnaires (QLQs) originally developed by the European Organization for Research and Treatment of Cancer (EORTC) [34,35]: the EORTC QLQ Core 30 (QLQ-C30), version 3, and the EORTC Breast Cancer-Specific QLQ (QLQ-BR23). These instruments have good test-retest reliability, high internal consistency in most scales, and show expected differences between patients in active chemotherapy and those in follow-up groups [36]. The QLQ-C30 incorporates five functional scales—physical functioning, role functioning, emotional functioning, cognitive functioning, and social functioning—nine symptom scales—fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties—as well as global health status and QoL scales [34]. The QLQ-BR23 is the module for breast cancer. It incorporates four functional scales—body image, sexual functioning, sexual enjoyment, and future perspective— and four symptom scales—systemic therapy side effects, breast symptoms, arm symptoms, and upset by hair loss [37]. According to the EORTC QLQ scoring manual, a high score for a functional scale represents a high or healthy level of functioning when using the QLQ-C30 and QLQ-BR23, a high score for global health status or QoL represents a high QoL when using the QLQ-C30, but a high score for a symptom scale or item represents a high level of symptomatology or problems when using the QLQ-C30 and QLQ-BR23 [38]. In this study, the instrument used was a paper-based or Web-based form. Participants could choose one of the evaluation form types according to their preferences (eg, desire to go paperless or efficiency when self-reporting [39]) at the two follow-up evaluations. The average time required to complete any of the evaluation forms was approximately 15 minutes, and most patients required no assistance. Following data collection, the internal consistency reliability of both instrument types used by the two groups was assessed: the Cronbach alpha levels at baseline, 1.5-month follow-up, and 3-month follow-up were .90-.96 using the QLQ-C30 and .71-.95 using the QLQ-BR23. The results indicated that both types of forms—paper based and Web based—had adequate reliability [40].

Data Analysis

All completed questionnaires were coded and analyzed using SPSS Statistics for Windows, version 20.0 (IBM Corp). Each item of the QLQ-C30 and QLQ-BR23 underwent linear transformation to obtain a score of 0-100 in accordance with the EORTC QLQ scoring manual [38]. For data values that were missing at random (MAR), the multi-imputation technique was used [38,41] and the average score of the observed data at the same follow-up interval in the same group was adopted to impute the MAR. The dummy variables—summary scores of the QLQ-C30 and QLQ-BR23—were adopted to represent the overall QoL outcome of each respondent and the formula was the sum of scores from each functional item scale and each revised symptom item (eg, the difference between 100 and each symptom item scale score). A higher summary score represents a higher QoL. Frequency and percentage as well as mean and standard deviation were used for descriptive statistics for clinical variables, the QLQ-C30, and the QLQ-BR23. We compared the baseline results of the control and experimental groups using chi-square tests for categorical variables and *t* tests for continuous variables. All analyses were intention-to-treat using a repeated-measures analysis and the generalized estimating equation (GEE) [42]. This method was used to account for the lack of adherence values over time and to detect any time × group effects among the target indicators. Significance was defined as a *P* value less than .05.

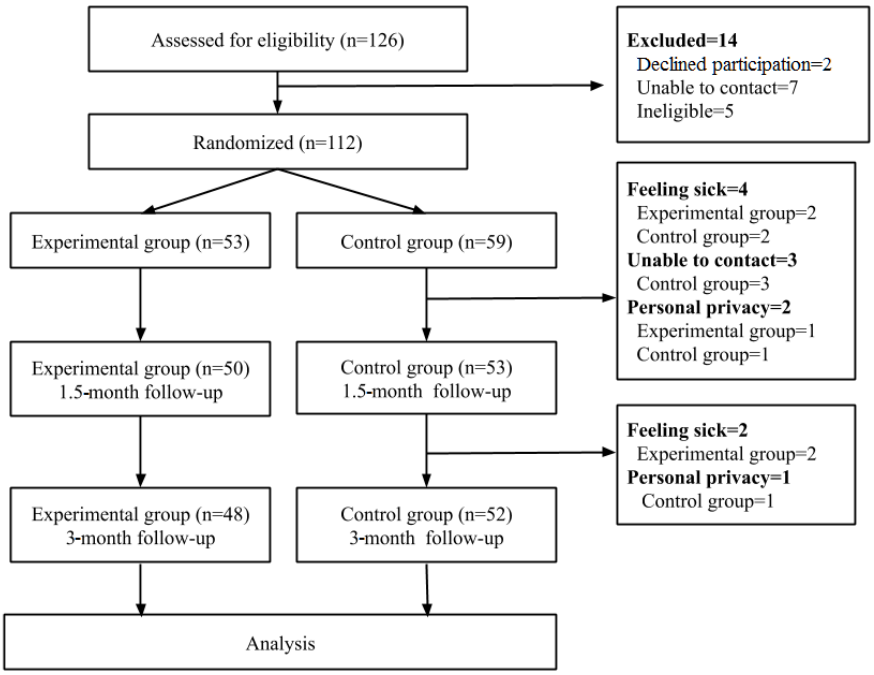
Results

Randomization and Attrition

Randomization and attrition data were organized according to the Consolidated Standards Of Reporting Trials (CONSORT) guidelines [43] (see Figure 1). A total of 112 eligible women, after initial diagnosis of nonmetastatic breast cancer, were enrolled in the study and randomly allocated to the experimental (n=53) or control (n=59) group. Reasons for patient dropout included feeling sick during treatment (n=6), lost to follow-up (n=3), and concern for their personal privacy (n=3). A total of 48 participants remained in the experimental group and 52 remained in the control group at the 3-month follow-up; the follow-up completion rate was 89.3% (100/112). The statistical power of 100 participants was 0.99, which is adequate according to Cohen [44].

At baseline, all participants were assessed using paper-based instruments. At the 1.5-month follow-up evaluation, half of the participants were assessed using Web-based instruments: 48 used paper-based forms versus 55 who used Web-based forms. At the 3-month follow-up evaluation, most participants were assessed using Web-based instruments: 22 used paper-based forms versus 78 who used Web-based forms. The internal consistency reliability of the two groups at each follow-up were adequate; the Cronbach alpha levels are shown in the Instruments section.

Figure 1. Study flowchart.



Participant Demographics

The demographics and baseline QoL from the two groups were similar. The largest age group was 50-64 years (51/112, 45.5%). Participants had bachelor's degrees or equivalent levels of education (34/112, 30.4%). Over half (78/112, 69.6%) of the participants were married. One-third of participants had two children (43/112, 38.4%). Close to half (52/112, 46.4%) of the participants were employed. The majority of participants did not have any comorbid diagnoses (98/112, 87.5%). The largest number of participants had stage II breast cancer (44/112, 39.3%) and received breast conservation surgery (76/112, 67.9%), chemotherapy (73/112, 65.2%), radiotherapy (74/112, 66.1%), and hormone therapy (60/112, 53.6%). There were no statistical differences between the control and experimental group participants with regard to any of the baseline characteristics. Detailed statistical results are shown in Table 2.

At the 1.5-month and 3-month follow-up evaluations, there were also no statistical differences between control and experimental group participants with respect to receiving chemotherapy (at 1.5 months, P=.40; at 3 months, P=.76), radiotherapy (at 1.5 months, P=.92; at 3 months, P=.31), and hormone therapy (at 1.5 months, P=.13; at 3 months, P=.06), which also showed homogeneity in the participants' medical treatment statuses at the two follow-up evaluations.

Table 2. Demographic characteristics and treatments of the participants.

Characteristic and treatment	Experimental group (n=53), n (%)	Control group (n=59), n (%)	All participants (N=112), n (%)	*P* value
Age in years				.33
20-34	8 (15)	4 (7)	12 (10.7)	
35-49	22 (42)	27 (46)	49 (43.8)	
50-64	23 (43)	28 (47)	51 (45.5)	
Education				.05
Junior high school	4 (8)	12 (20)	16 (14.3)	
Senior high school	14 (26)	14 (24)	28 (25.0)	
College	9 (17)	13 (22)	22 (19.6)	
Bachelor or equivalent	18 (34)	16 (27)	34 (30.4)	
Master or equivalent	8 (15)	4 (7)	12 (10.7)	
Marital status				.21
Unmarried	13 (25)	10 (17)	23 (20.5)	
Married	37 (70)	41 (69)	78 (69.6)	
Divorced	1 (2)	6 (10)	7 (6.3)	
Widowed	2 (4)	2 (3)	4 (3.6)	
Number of children				.28
0	16 (30)	15 (25)	31 (27.7)	
1	7 (13)	7 (12)	14 (12.5)	
2	22 (42)	21 (36)	43 (38.4)	
More than 2	8 (15)	16 (27)	24 (21.4)	
Work status				.79
Unemployed	6 (11)	8 (14)	14 (12.5)	
Housewife	18 (34)	20 (34)	38 (33.9)	
Retired	4 (8)	4 (7)	8 (7.1)	
Employed	25 (47)	27 (46)	52 (46.4)	
Past disease history				.18
No	44 (83)	54 (92)	98 (87.5)	
Yes	9 (17)	5 (8)	14 (12.5)	
Breast cancer stage				.55
0	5 (9)	4 (7)	9 (8.0)	
I	22 (42)	20 (34)	42 (37.5)	
II	17 (32)	27 (46)	44 (39.3)	
III	9 (17)	8 (14)	17 (15.2)	
Surgery				.55
Conservation therapy	38 (72)	38 (64)	76 (67.9)	
Mastectomy	15 (28)	21 (36)	36 (32.1)	
Chemotherapy				.32
No	21 (40)	18 (31)	39 (34.8)	
Yes	32 (60)	41 (69)	73 (65.2)	
Radiotherapy				.11
No	14 (26)	24 (41)	38 (33.9)	
Yes	39 (74)	35 (59)	74 (66.1)	

Characteristic and treatment	Experimental group (n=53), n (%)	Control group (n=59), n (%)	All participants (N=112), n (%)	*P* value
Hormone therapy				.08
No	20 (38)	32 (54)	52 (46.4)	
Yes	33 (62)	27 (46)	60 (53.6)	

Changes in Quality-of-Life Questionnaire Core 30 Indicators

Before instruction (T0), the mean total summary scores of the QLQ-C30 for the experimental and control groups were 74.47 (SD 14.96) and 78.30 (SD 12.59), respectively, with no difference between the groups (*P*=.14).

At the 1.5-month and 3-month follow-up evaluations, these scores were 79.13 (SD 15.31) and 83.45 (SD 10.85) in the experimental group, respectively, and 79.49 (SD 12.41) and 82.23 (SD 12.07) in the control group, respectively. The mean summary scores for the QLQ-C30 in both groups showed a significant improvement by 3 months: the difference between the T2 and T0 scores was 8.98 in the experimental group and 3.93 in the control group. GEE analysis showed a statistically significant difference (*P*=.03) in the interaction between groups and time, which meant the experimental group had a greater mean summary score for the QLQ-C30 (difference=5.05) than the control group after the intervention at the 3-month assessment (see Table 3).

Table 3. Generalized estimating equation analysis of longitudinal outcome of the Quality-of-Life Questionnaire Core 30 (N=112).

Step	Functional scales[a], mean			*P* value	Symptom scales[b], mean			*P* value	Global health status or quality of life[a], mean			*P* value	Total summary scores[a], mean			*P* value
	Exp[c]	Con[d]	Diff[e]		Exp	Con	Diff		Exp	Con	Diff		Exp	Con	Diff	
T0[f]	73.59	77.18	-3.58	.25	24.71	20.99	3.73	.16	57.08	63.28	-6.20	.08	74.47	78.30	-3.82	.14
T1[g]	79.64	78.63	1.01	.71	21.14	19.50	1.64	.52	68.33	70.75	-2.42	.51	79.13	79.49	-0.37	.89
T2[h]	82.76	80.64	2.12	.38	16.10	16.62	-0.52	.79	73.44	74.36	-0.92	.76	83.45	82.23	1.22	.55
T1-T0	6.05	1.46	4.59	.06	-3.58	-1.49	-2.09	.32	11.26	7.48	3.78	.38	4.65	1.19	3.46	.09
T2-T0	9.17	3.47	5.71	.04	-8.61	-4.37	-4.25	.06	16.36	11.08	5.28	.18	8.98	3.93	5.04	.03

[a]Higher scores correspond with better quality of life.

[b]Lower scores correspond with better quality of life.

[c]Exp: experimental group.

[d]Con: control group.

[e]Diff: difference between the experimental and control groups.

[f]T0: baseline or pretest evaluation.

[g]T1: 1.5-month follow-up evaluation.

[h]T2: 3-month follow-up evaluation.

Changes in Breast Cancer-Specific Quality-of-Life Questionnaire Indicators

Before instruction (T0), the QLQ-BR23 mean summary scores of the experimental and control groups were 59.68 (SD 12.46) and 61.21 (SD 12.65), respectively, with no difference between the groups (*P*=.52).

At 1.5-month and 3-month follow-up evaluations, these scores were 62.56 (SD 13.10) and 65.53 (SD 10.29) in the experimental group, respectively, and 62.13 (SD 12.74) and 63.13 (SD 12.66) in the control group, respectively. The mean summary scores for QLQ-BR23 in both groups increased significantly by 3 months: the difference between the T2 and T0 scores was 5.85 in the experimental group and 1.92 in the control group. GEE analysis showed a statistically significant difference (*P*=.04) in the interaction between groups and time, which meant the experimental group had a greater total summary score for the QLQ-BR23 (difference=3.93) than the control group after the intervention at the 3-month assessment (see Table 4).

Table 4. Generalized estimating equation analysis of longitudinal outcome of the Breast Cancer-Specific Quality-of-Life Questionnaire (N=112).

Step	Functional scales[a], mean			*P* value	Symptom scales[b], mean			*P* value	Total summary scores[a], mean			*P* value
	Exp[c]	Con[d]	Diff[e]		Exp	Con	Diff		Exp	Con	Diff	
T0[f]	46.54	47.32	-0.78	.76	27.38	25.45	1.93	.51	59.68	61.20	-1.53	.52
T1[g]	49.39	47.06	2.32	.35	25.11	24.08	1.03	.72	62.56	62.13	0.43	.86
T2[h]	34.65	32.82	1.83	.22	21.69	23.46	-1.78	.51	65.53	63.13	2.40	.24
T1-T0	2.85	-0.25	3.10	.17	-2.27	-1.37	-0.90	.71	2.88	0.93	1.95	.30
T2-T0	-11.89	-14.49	2.61	.31	-5.69	-1.99	-3.70	.14	5.85	1.92	3.93	.04

[a]Higher scores correspond with better quality of life.

[b]Lower scores correspond with better quality of life.

[c]Exp: experimental group.

[d]Con: control group.

[e]Diff: difference between the experimental and control groups.

[f]T0: baseline or pretest evaluation.

[g]T1: 1.5-month follow-up evaluation.

[h]T2: 3-month follow-up evaluation.

Sensitivity Analysis

We performed a sensitivity analysis for the attrition cases using GEE analysis. The results showed that the mean summary scores also increased in the QLQ-C30 (difference between experimental and control groups was 5.05, *P*=.07) and the QLQ-BR23 (difference between experimental and control groups was 3.93, *P*=.07), with no statistically significant difference between the two groups noted at the 3-month assessment (see Table 5).

Table 5. Generalized estimating equation analysis of longitudinal outcome for the Quality-of-Life Questionnaire Core 30 (QLQ-C30) and the Breast Cancer-Specific Quality-of-Life Questionnaire (QLQ-BR23) for sensitivity analysis.

Step	QLQ-C30 mean summary scores			*P* value	QLQ-BR23 mean summary scores			*P* value
	Exp[a]	Con[b]	Diff[c]		Exp	Con	Diff	
T0[d]	74.47	78.30	-3.83	.14	59.68	61.21	-1.53	.52
T1[e]	79.27	79.49	-0.22	.89	63.04	62.13	0.91	.87
T2[f]	83.45	82.23	1.22	.99	65.53	63.13	2.40	.40
T1-T0	4.80	1.19	3.61	.11	3.36	0.92	2.44	.34
T2-T0	8.98	3.93	5.05	.07	5.85	1.92	3.93	.07

[a]Exp: experimental group.

[b]Con: control group.

[c]Diff: difference between the experimental and control groups.

[d]T0: baseline or pretest evaluation.

[e]T1: 1.5-month follow-up evaluation.

[f]T2: 3-month follow-up evaluation.

Discussion

Principal Findings

Our team developed the BCSMS app with eight main features, in line with the results from our previous study, to support self-management for women after a first diagnosis of breast cancer. After 3 months of repeated QoL evaluation, the principal results showed that for the QLQ-C30, both groups had improvement in their functional scales (9.17 for the experimental group vs 3.47 for the control group), symptom scales (-8.61 for the experimental group vs -4.37 for the control group), global health status (16.36 for the experimental group vs 11.08 for the control group), and total summary scores (8.98 for the experimental group vs 3.93 for the control group). The BCSMS app user group participants had significant improvement in their functional scores (difference=5.71, *P*=.04) and total summary scores (difference=5.04, *P*=.03) compared to the control group.

Based on the results of the QLQ-BR23, neither group had improvement in functional scales (-11.98 for the experimental group vs -14.49 for the control group) by the third month. However, improvement was noted in symptom scales (-5.69 for the experimental group vs -1.99 for the control group) and total summary scores (5.85 for the experimental group vs 1.92 for the control group) by the third month. The BCSMS app user

group had a significant improvement in total summary scores (difference=3.93, *P*=.04) when compared to the control group.

After sensitivity analysis (ie, no imputation for attrition cases), the total summary scores for the QLQ-C30 (difference=5.05, *P*=.07) and the QLQ BR23 (difference=3.93, *P*=.07) also showed improvement by the third month between the two groups but it was not significant. Based on these findings, further evaluation of QoL among women with a first diagnosis of breast cancer after using the BCSMS app is warranted.

Comparison With Prior Work

The overall QLQ-C30 and QLQ-BR23 scores (see Tables 3 and 4) in both groups in this study were close to those from a multicenter, cross-sectional study in Taiwan [19]. They were also similar to the QoL scores for women after 1 year following a breast cancer diagnosis in a 10-year, long-term follow-up study in Germany [45]. This may be due to standardized medical treatments for each stage of breast cancer that were adopted regionally and globally [3-5]. Mobile health apps can effectively support patients with chronic disease self-management [46-48]. Patients with a first diagnosis of breast cancer need knowledge about the disease and to learn about self-management in order to live well with the cancer (eg, surgery-wound pain relief, lymphedema prevention, and controlling body weight). Others have demonstrated that mobile health apps are associated with improved knowledge of the disease and self-management, including self-efficacy with performing shoulder exercises during and after treatment, symptom relief, and QoL during chemotherapy [49-51]. In our study, the BCSMS app also demonstrated improved QoL for Taiwanese women with breast cancer. Among the eight main features included in the BCSMS app, which was developed with the user-centered approach [52], was a feature that provided evidence or *knowledge* about breast cancer (eg, cancer stages, treatments, side effects, lymphedema prevention, and relapse) and addressed basic information for the patients. According to the prior study, higher levels of anxiety were experienced by preoperative breast cancer patients who received information delivered via a mobile app than by patients who did not use the app [53]. In our study, every subject was postoperative; we did not know the BCSMS app would increase patients' anxiety levels when they had *knowledge* about breast cancer before operation. Our results showed that the group who used the BCSMS app had higher scores in functional scales—physical functioning, role functioning, emotional functioning, cognitive functioning, and social functioning—than did the group who did not use the app at the 3-month follow-up. This might be because the BCSMS app not only provided the evidence or *knowledge* about breast cancer but also provided *emotional support* features to prevent anxiety and depression, including information on mental support, music therapy, mindfulness activities, sleeping well, and acupuncture point massage; *social resource* information, including information on nonprofit supporting organizations, financial support, and housekeeping; *experience sharing*, including encouraging information from other senior survivors and health care professionals; and *expert consulting*, including over 100 frequently asked questions. With this positive and useful information from the BCSMS app, patients were supported in overcoming cancer-related emotional disturbances.

In addition, the BCSMS app provided videos that demonstrated effective body movements to prevent lymphedema and low-intensity exercises in the *exercise and rehabilitation* feature. This feature may have reduced patients' symptom scale scores, including for fatigue, pain, dyspnea, insomnia, appetite loss, and constipation. Lowering of these scores benefitted symptom control during the first year of cancer treatment (eg, cancer-related fatigue improvement) among experimental subjects and had positive outcomes, similar to those seen in prior studies [54,55].

In the *personal health record* feature for tracking treatments and side effects, subjects in the experimental group were able to record details about their personal medical treatments (eg, date of surgery, period of radiotherapy, medication for chemotherapy, and hormone therapy), record physical-related self-measurements (eg, body temperature and arm circumference), view data graphs, and receive abnormal-data warnings (eg, reminders to patients to consult their psychologists when their emotions were self-assessed as poor through the Brief Symptom Rating Scale) (publication is forthcoming). This feature was also similar to the patient-reported outcomes for symptom monitoring during routine cancer treatment; the results showed that integration of patient-reported outcomes into the routine care of patients with metastatic cancer was associated with increased survival compared with usual care [56]. Having all of the features integrated into one mobile app, as well as the app's accessibility and ease of use regarding breast cancer self-management, may have contributed to the improvement in the patients' QoL.

Limitations

There were two main limitations of this study. The first was maturation bias, as our study followed participants for only 3 months. Although the control group did not use the BCSMS app, they still had supportive care from health care professionals (eg, physicians and oncology case managers), and perhaps their information-searching competence helped them with self-management and to increase their QoL. The second was that we did not know the actual frequency of use of the BCSMS app among the experimental group. To prevent the Hawthorne effect (ie, causing nonroutine behavior as a result of being observed) [57], the research team did not remind the subjects in the experimental group to use BCSMS app. Such limitations might have influenced the results in this study.

Conclusions

The purpose of the study was to investigate the QoL of women with breast cancer in Taiwan after using the BCSMS app. According to the results, women with a first diagnosis of breast cancer appear to have experienced increased QoL after receiving cancer treatment. The BCSMS app provided users with supportive evidence to promote their QoL, compared with those who did not use the app. This might be because the BCSMS app delivered comprehensive information (eg, eight main features developed during our prior work) to women with breast cancer when they were first diagnosed and the app and its features supported their needs during their treatment. Next steps would be to introduce the BCSMS app to more breast cancer

patients and to further evaluate whether QoL is sustained among these patients.

Implications

Using objective data, such as users' log-in frequencies, most-used features, and content analysis from the *personal health record* for treatment and side effects, could be investigated to better understand how the BCSMS app helps women with breast cancer. In addition, subjective feedback from BCSMS app users is important for future modifications and enhancements of the app. In the future, the BCSMS app could be introduced to breast cancer patients in different countries after culturally sensitive content is added (eg, *diet and nutrition* for breast cancer patients and *social resource* information) and when multiple languages can be incorporated into the interface (eg, English).

Acknowledgments

We appreciate the funding from the Ministry of Science and Technology (MOST107 - 2314 - B - 010 - 052) and we acknowledge all the team members who participated in this research as well as the Taiwan Breast Cancer Foundation. We would like to give special thanks to Jhou-Liang Lian, Kong-Hao Chen, and Zih-Bin Chen for their support.

Authors' Contributions

ICH and HYL contributed to the study design and implementation, data analysis, interpretation of the findings, and preparation of the manuscript. SHS contributed to the technology support. KJC, HCT, and AJT contributed to the study implementation. PCD contributed to the editing and revision of the manuscript in English.

Conflicts of Interest

None declared.

This randomized study was only retrospectively registered, explained by authors with "We did not prospectively register this trial due to a lack of awareness or understanding of this requirement on our first Randomized Control Trial study submission". The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low and the study was considered formative, guiding the development of the application [or other reasons for the exception, as argued by the authors]. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 384 KB - mhealth_v8i3e17084_app1.pdf]

References

1. Health Promotion Administration. Cancer Registry Annual Report, 2015. Taipei, Taiwan: Ministry of Health and Welfare; 2017. URL: https://www.hpa.gov.tw/Pages/ashx/File.ashx?FilePath=~/File/Attach/8084/File_7635.pdf [accessed 2019-10-08]
2. Health Promotion Administration. Cancer Registry Annual Report, 2016. Taipei, Taiwan: Ministration of Health and Welfare; 2018 Dec. URL: https://www.hpa.gov.tw/Pages/ashx/File.ashx?FilePath=~/File/Attach/10227/File_11644.pdf [accessed 2019-12-30]
3. Lin CH. Chemotherapy in breast cancer. Formos J Med 2013;17(4):404-411. [doi: 10.6320/fjm.2013.17(4).08]
4. Matsen CB, Neumayer LA. Breast cancer: A review for the general surgeon. JAMA Surg 2013 Oct;148(10):971-979. [doi: 10.1001/jamasurg.2013.3393] [Medline: 23986370]
5. Diaby V, Tawk R, Sanogo V, Xiao H, Montero AJ. A review of systematic reviews of the cost-effectiveness of hormone therapy, chemotherapy, and targeted therapy for breast cancer. Breast Cancer Res Treat 2015 May;151(1):27-40 [FREE Full text] [doi: 10.1007/s10549-015-3383-6] [Medline: 25893588]
6. Lo C, Liao FZ, Chen SL. Symptom experiences of breast cancer patients. J Nurs 2017;64(2):19-27. [doi: 10.6224/jn.000021]
7. Browall M, Brandberg Y, Nasic S, Rydberg P, Bergh J, Rydén A, et al. A prospective exploration of symptom burden clusters in women with breast cancer during chemotherapy treatment. Support Care Cancer 2017 May;25(5):1423-1429 [FREE Full text] [doi: 10.1007/s00520-016-3527-1] [Medline: 27981366]
8. Hermelink K, Bühner M, Sckopke P, Neufeld F, Kaste J, Voigt V, et al. Chemotherapy and post-traumatic stress in the causation of cognitive dysfunction in breast cancer patients. J Natl Cancer Inst 2017 Oct 01;109(10):1-15. [doi: 10.1093/jnci/djx057] [Medline: 28521364]
9. Wiśniewska I, Jochymek B, Lenart-Lipińska M, Chabowski M. The pharmacological and hormonal therapy of hot flushes in breast cancer survivors. Breast Cancer 2016 Mar;23(2):178-182 [FREE Full text] [doi: 10.1007/s12282-015-0655-2] [Medline: 26498637]

10. Savard J, Ivers H, Savard M, Morin CM. Cancer treatments and their side effects are associated with aggravation of insomnia: Results of a longitudinal study. Cancer 2015 May 15;121(10):1703-1711 [FREE Full text] [doi: 10.1002/cncr.29244] [Medline: 25677509]
11. Jung MS, Cimprich B. Cognitive deficits in Korean women treated with chemotherapy for breast cancer. Cancer Nurs 2014;37(3):E31-E42. [doi: 10.1097/ncc.0b013e3182980383]
12. Yee C, Wang K, Asthana R, Drost L, Lam H, Lee J, et al. Radiation-induced skin toxicity in breast cancer patients: A systematic review of randomized trials. Clin Breast Cancer 2018 Oct;18(5):e825-e840. [doi: 10.1016/j.clbc.2018.06.015] [Medline: 30072193]
13. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: A review. Patient Educ Couns 2002 Oct;48(2):177-187. [doi: 10.1016/s0738-3991(02)00032-0]
14. McCorkle R, Ercolano E, Lazenby M, Schulman-Green D, Schilling LS, Lorig K, et al. Self-management: Enabling and empowering patients living with cancer as a chronic illness. CA Cancer J Clin 2011;61(1):50-62 [FREE Full text] [doi: 10.3322/caac.20093] [Medline: 21205833]
15. Li MF, Pan LC. Applications on mobile health. J Long Term Care 2012;16(3):237-250.
16. Kuo CC, Wang HH. Applying self-management intervention studies to patients with chronic illness: A systematic review. J Nurs Healthc Res 2012;8(1):50-60. [doi: 10.6225/jnhr.8.1.52]
17. Loh SY, Packer T, Chinna K, Quek KF. Effectiveness of a patient self-management programme for breast cancer as a chronic illness: A non-randomised controlled clinical trial. J Cancer Surviv 2013 Sep;7(3):331-342. [doi: 10.1007/s11764-013-0274-x] [Medline: 23519654]
18. World Health Organization. 2019 Dec. WHOQOL: Measuring quality of life URL: http://www.who.int/healthinfo/survey/whoqol-qualityoflife/en/ [accessed 2019-12-12]
19. Tsai H, Kuo NC, Chung KP. Quality of life of breast cancer survivors following breast-conserving therapy versus mastectomy: A multicenter study in Taiwan. Jpn J Clin Oncol 2017 Oct 01;47(10):909-918. [doi: 10.1093/jjco/hyx099] [Medline: 28981734]
20. Fanakidou I, Zyga S, Alikari V, Tsironi M, Stathoulis J, Theofilou P. Mental health, loneliness, and illness perception outcomes in quality of life among young breast cancer patients after mastectomy: The role of breast reconstruction. Qual Life Res 2018 Feb;27(2):539-543. [doi: 10.1007/s11136-017-1735-x] [Medline: 29119452]
21. Hamer J, McDonald R, Zhang L, Verma S, Leahey A, Ecclestone C, et al. Quality of life (QOL) and symptom burden (SB) in patients with breast cancer. Support Care Cancer 2017 Feb;25(2):409-419. [doi: 10.1007/s00520-016-3417-6] [Medline: 27696078]
22. Kessel KA, Vogel MM, Kessel C, Bier H, Biedermann T, Friess H, et al. Mobile health in oncology: A patient survey about app-assisted cancer care. JMIR Mhealth Uhealth 2017 Jun 14;5(6):e81 [FREE Full text] [doi: 10.2196/mhealth.7689] [Medline: 28615159]
23. Kessel KA, Vogel MM, Schmidt-Graf F, Combs SE. Mobile apps in oncology: A survey on health care professionals' attitude toward telemedicine, mHealth, and oncological apps. J Med Internet Res 2016 Nov 24;18(11):e312 [FREE Full text] [doi: 10.2196/jmir.6399] [Medline: 27884810]
24. Zhu J, Ebert L, Guo D, Yang S, Han Q, Chan SW. Mobile breast cancer e-support program for Chinese women with breast cancer undergoing chemotherapy (Part 1): Qualitative study of women's perceptions. JMIR Mhealth Uhealth 2018 Apr 11;6(4):e85 [FREE Full text] [doi: 10.2196/mhealth.9311] [Medline: 29643056]
25. Rincon E, Monteiro-Guerra F, Rivera-Romero O, Dorronzoro-Zubiete E, Sanchez-Bocanegra CL, Gabarron E. Mobile phone apps for quality of life and well-being assessment in breast and prostate cancer patients: Systematic review. JMIR Mhealth Uhealth 2017 Dec 04;5(12):e187 [FREE Full text] [doi: 10.2196/mhealth.8741] [Medline: 29203459]
26. Giunti G, Giunta D, Guisado-Fernandez E, Bender J, Fernandez-Luque L. A biopsy of breast cancer mobile applications: State of the practice review. Int J Med Inform 2018 Feb;110:1-9 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.022] [Medline: 29331247]
27. Lee H, Uhm KE, Cheong IY, Yoo JS, Chung SH, Park YH, et al. Patient satisfaction with mobile health (mHealth) application for exercise intervention in breast cancer survivors. J Med Syst 2018 Nov 06;42(12):254. [doi: 10.1007/s10916-018-1096-1] [Medline: 30402781]
28. World Medical Association. World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
29. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
30. Cohen J. The analysis of variance and covariance. In: Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988:273-406.
31. Day SJ, Altman DG. Statistics notes: Blinding in clinical trials and other studies. BMJ 2000;321(7259):504 [FREE Full text] [doi: 10.1136/bmj.321.7259.504] [Medline: 10948038]
32. Persaud N, Heneghan C, Catalogue of Bias Collaboration. Catalogue of Bias. Novelty bias URL: https://catalogofbias.org/biases/novelty-bias/ [accessed 2019-12-21]
33. Randomization.com. URL: http://www.randomization.com [accessed 2018-12-01]

34. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 03;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
35. EORTC Quality of Life. Questionnaires: Breast URL: https://qol.eortc.org/questionnaire/qlq-br23/ [accessed 2019-08-02]
36. Chie W, Chang K, Huang C, Kuo W. Quality of life of breast cancer patients in Taiwan: Validation of the Taiwan Chinese version of the EORTC QLQ-C30 and EORTC QLQ-BR23. Psychooncology 2003;12(7):729-735. [doi: 10.1002/pon.727] [Medline: 14502597]
37. Sprangers MA, Groenvold M, Arraras JI, Franklin J, te Velde A, Muller M, et al. The European Organization for Research and Treatment of Cancer breast cancer-specific quality-of-life questionnaire module: First results from a three-country field study. J Clin Oncol 1996 Oct;14(10):2756-2768. [doi: 10.1200/JCO.1996.14.10.2756] [Medline: 8874337]
38. Fayers P, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A. EORTC QLQ-C30 Scoring Manual. 3rd edition. Brussels, Belgium: European Organisation for Research and Treatment of Cancer; 2001.
39. Noyes JM, Garland KJ. Computer- vs paper-based tasks: Are they equivalent? Ergonomics 2008 Sep;51(9):1352-1375. [doi: 10.1080/00140130802170387] [Medline: 18802819]
40. Henson RK. Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. Meas Eval Couns Dev 2019 Aug 29;34(3):177-189. [doi: 10.1080/07481756.2002.12069034]
41. Fitzmaurice GM, Laird NM, Ware JH. Missing data and dropout: Multiple imputation and weighting methods. In: Applied Longitudinal Analysis. 2nd edition. Hoboken, NJ: John Wiley & Sons; 2011:515-550.
42. Twisk J, de Vente W. Attrition in longitudinal studies. J Clin Epidemiol 2002 Apr;55(4):329-337 [FREE Full text] [doi: 10.1016/s0895-4356(01)00476-0]
43. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. BMC Med 2010 Mar 24;8:18 [FREE Full text] [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
44. Cohen J. The concepts of power analysis. In: Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988:1-17.
45. Koch L, Jansen L, Herrmann A, Stegmaier C, Holleczek B, Singer S, et al. Quality of life in long-term breast cancer survivors: A 10-year longitudinal population-based study. Acta Oncol 2013 Mar 20;52(6):1119-1128. [doi: 10.3109/0284186x.2013.774461]
46. Farzandipour M, Nabovati E, Sharif R, Arani MH, Anvari S. Patient self-management of asthma using mobile health applications: A systematic review of the functionalities and effects. Appl Clin Inform 2017 Oct;8(4):1068-1081 [FREE Full text] [doi: 10.4338/ACI-2017-07-R-0116] [Medline: 29241254]
47. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: Systematic review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
48. Duke DC, Barry S, Wagner DV, Speight J, Choudhary P, Harris MA. Distal technologies and type 1 diabetes management. Lancet Diabetes Endocrinol 2018 Feb;6(2):143-156. [doi: 10.1016/S2213-8587(17)30260-7] [Medline: 28867311]
49. Zhu J, Ebert L, Liu X, Wei D, Chan SW. Mobile breast cancer e-support program for Chinese women with breast cancer undergoing chemotherapy (Part 2): Multicenter randomized controlled trial. JMIR Mhealth Uhealth 2018 Apr 30;6(4):e104 [FREE Full text] [doi: 10.2196/mhealth.9438] [Medline: 29712622]
50. Harder H, Holroyd P, Burkinshaw L, Watten P, Zammit C, Harris PR, et al. A user-centred approach to developing bWell, a mobile app for arm and shoulder exercises after breast cancer treatment. J Cancer Surviv 2017 Dec;11(6):732-742 [FREE Full text] [doi: 10.1007/s11764-017-0630-3] [Medline: 28741202]
51. Cruz FOAM, Vilela RA, Ferreira EB, Melo NS, Reis PEDD. Evidence on the use of mobile apps during the treatment of breast cancer: Systematic review. JMIR Mhealth Uhealth 2019 Aug 27;7(8):e13245 [FREE Full text] [doi: 10.2196/13245] [Medline: 31456578]
52. Hou IC, Lan MF, Shen SH, Tsai PY, Chang KJ, Tai HC, et al. Developing a breast cancer self-management support mHealth application in Taiwan: A design thinking approach [In press]. JMIR mHealth and uHealth 2020 (forthcoming). [doi: 10.2196/15780]
53. Foley N, O'Connell E, Lehane E, Livingstone V, Maher B, Kaimkhani S, et al. PATI: Patient accessed tailored information: A pilot study to evaluate the effect on preoperative breast cancer patients of information delivered via a mobile application. Breast 2016 Dec;30:54-58. [doi: 10.1016/j.breast.2016.08.012] [Medline: 27611236]
54. van Vulpen JK, Peeters PH, Velthuis MJ, van der Wall E, May AM. Effects of physical exercise during adjuvant breast cancer treatment on physical and psychosocial dimensions of cancer-related fatigue: A meta-analysis. Maturitas 2016 Mar;85:104-111. [doi: 10.1016/j.maturitas.2015.12.007] [Medline: 26857888]
55. van Waart H, Stuiver MM, van Harten WH, Geleijn E, Kieffer JM, Buffart LM, et al. Effect of low-intensity physical activity and moderate- to high-intensity physical exercise during adjuvant chemotherapy on physical fitness, fatigue, and chemotherapy completion rates: Results of the PACES randomized clinical trial. J Clin Oncol 2015 Jun 10;33(17):1918-1927. [doi: 10.1200/JCO.2014.59.1081] [Medline: 25918291]
56. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]

57. McCarney R, Warner J, Iliffe S, van Haselen R, Griffin M, Fisher P. The Hawthorne Effect: A randomised, controlled trial. BMC Med Res Methodol 2007 Jul 03;7:30 [FREE Full text] [doi: 10.1186/1471-2288-7-30] [Medline: 17608932]

Abbreviations

BCSMS: breast cancer self-management support
CONSORT: Consolidated Standards Of Reporting Trials
EORTC: European Organization for Research and Treatment of Cancer
GEE: generalized estimating equation
IRB: Institutional Review Board
MAR: missing at random
QLQ: Quality-of-Life Questionnaire
QLQ-BR23: Breast Cancer-Specific Quality-of-Life Questionnaire
QLQ-C30: Quality-of-Life Questionnaire Core 30
QoL: quality of life
T0: pretest evaluation
T1: 1.5-month follow-up evaluation
T2: 3-month follow-up evaluation

Edited by G Eysenbach; submitted 16.11.19; peer-reviewed by E Rincon, K Crew, Z Zhou; comments to author 12.12.19; revised version received 25.12.19; accepted 26.01.20; published 04.03.20.

Please cite as:
Hou IC, Lin HY, Shen SH, Chang KJ, Tai HC, Tsai AJ, Dykes PC
Quality of Life of Women After a First Diagnosis of Breast Cancer Using a Self-Management Support mHealth App in Taiwan: Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(3):e17084
URL: http://mhealth.jmir.org/2020/3/e17084/
doi:10.2196/17084
PMID:32130181

Original Paper

Incorporating Behavioral Trigger Messages Into a Mobile Health App for Chronic Disease Management: Randomized Clinical Feasibility Trial in Diabetes

Scott Sittig[1], BS, MHI, PhD; Jing Wang[2], BSN, MPH, PhD; Sriram Iyengar[3], BS, PhD; Sahiti Myneni[4], BS, PhD; Amy Franklin[4], BS, PhD

[1]School of Computing, University of South Alabama, Mobile, AL, United States

[2]School of Nursing, University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

[3]College of Medicine Phoenix, The University of Arizona, Phoenix, AZ, United States

[4]School of Biomedical Informatics, University of Texas Health Science Center Houston, Houston, TX, United States

Corresponding Author:
Scott Sittig, BS, MHI, PhD
School of Computing
University of South Alabama
150 Jaguar Drive
Mobile, AL, 36688
United States
Phone: 1 2514607576
Email: sittig@southalabama.edu

Abstract

Background: Although there is a rise in the use of mobile health (mHealth) tools to support chronic disease management, evidence derived from theory-driven design is lacking.

Objective: The objective of this study was to determine the impact of an mHealth app that incorporated theory-driven trigger messages. These messages took different forms following the Fogg behavior model (FBM) and targeted self-efficacy, knowledge, and self-care. We assess the feasibility of our app in modifying these behaviors in a pilot study involving individuals with diabetes.

Methods: The pilot randomized unblinded study comprised two cohorts recruited as employees from within a health care system. In total, 20 patients with type 2 diabetes were recruited for the study and a within-subjects design was utilized. Each participant interacted with an app called capABILITY. capABILITY and its affiliated trigger (text) messages integrate components from social cognitive theory (SCT), FBM, and persuasive technology into the interactive health communications framework. In this within-subjects design, participants interacted with the capABILITY app and received (or did not receive) text messages in alternative blocks. The capABILITY app alone was the control condition along with trigger messages including spark and facilitator messages. A repeated-measures analysis of variance (ANOVA) was used to compare adherence with behavioral measures and engagement with the mobile app across conditions. A paired sample *t* test was utilized on each health outcome to determine changes related to capABILITY intervention, as well as participants' classified usage of capABILITY.

Results: Pre- and postintervention results indicated statistical significance on 3 of the 7 health survey measures (general diet: *P*=.03; exercise: *P*=.005; and blood glucose: *P*=.02). When only analyzing the high and midusers (n=14) of capABILITY, we found a statistically significant difference in both self-efficacy (*P*=.008) and exercise (*P*=.01). Although the ANOVA did not reveal any statistically significant differences across groups, there is a trend among spark conditions to respond more quickly (ie, shorter log-in lag) following the receipt of the message.

Conclusions: Our theory-driven mHealth app appears to be a feasible means of improving self-efficacy and health-related behaviors. Although our sample size is too small to draw conclusions about the differential impact of specific forms of trigger messages, our findings suggest that spark triggers may have the ability to cue engagement in mobile tools. This was demonstrated with the increased use of capABILITY at the beginning and conclusion of the study depending on spark timing. Our results suggest that theory-driven personalization of mobile tools is a viable form of intervention.

Trial Registration: ClinicalTrials.gov NCT04132089; http://clinicaltrials.gov/ct2/show/NCT004122089

(JMIR Mhealth Uhealth 2020;8(3):e15927) doi:10.2196/15927

KEYWORDS

mHealth; persuasive technology; Fogg behavior model; triggers; messages; interactive health communication application; self-efficacy; social cognitive theory; self-management; knowledge

Introduction

Background

The utilization of cell phones and, in particular, smartphones continues to rise. As of 2018, 95% of American adults own a cell phone, and 81% own a smartphone [1]. With the increase in smartphone utilization, we are seeing increasing use of mobile health (mHealth) wearable and sensing technology for patients with chronic diseases. By leveraging the prevalence of smartphones, we can create theoretically driven mHealth solutions that would engage patients in their chronic disease management while targeting sustainable behavior change.

Understanding how to engage patients (consumers) in their own behavior and health management, particularly as it relates to self-management for chronic conditions, is a daunting task. However, through the use of mHealth tools, we are able to design new techniques to promote patient engagement, which includes combining theoretical principles from behavior change and persuasive technology into existing mHealth design architectures [2-4]. Utilizing persuasive technology in which the patient interacts with an mHealth app while receiving trigger messages can promote user engagement, improve motivation, and bolster patients' belief in their own ability (self-efficacy) to manage their complex chronic health condition [5,6]. Self-efficacy refers to a person's belief that they can accomplish a task to produce a given outcome and has been shown to lead to positive behavior change and improved clinical outcomes, particularly in patients with chronic illnesses (ie, diabetes mellitus) [7-9]. In addition, mHealth apps can provide users with extensive educational material to improve self-efficacy and to simplify behavior change.

Behavioral trigger messages, or relevant text messages, are one way of facilitating behavior change by cueing targeted actions and providing reinforcement as needed (ie, increasing motivation and simplifying tasks to improve ability). Although shown to be useful for improving self-efficacy and self-management, most studies utilizing trigger messages have focused solely on reminder messages [10-12]. Expanding trigger messages to include other forms of engagement may also lead to positive effects on behavior change [13]. Recent studies have shown that the integration of behavior change theories into mHealth apps and digital health interventions can lead to effective designs in engaging the user and improving outcomes [14-16].

The utilization of interactive health communication applications (IHCA) frameworks have similarly been shown to be effective for chronic disease management (ie, type 2 diabetes) as it relates to knowledge and self-efficacy [3,17-19]. However, even with these critical breakthroughs in mHealth, there are still gaps in the development of mHealth apps for chronic disease management that focus on behavior change. These gaps include the following: embedding multiple theoretical constructs such as persuasive technology and behavior change theories into an mHealth system design, and the utilization of behavioral trigger messages instead of simple reminder messages for cueing specific behavioral tasks.

Mobile Health and Persuasive Technology

As of late 2017, 325,000 mHealth apps were available for download, with 78,000 new mHealth apps added in 2017 [20]. The volume of these programs reflects hope and interest in the ability of mHealth to transform health care [21-24]. Due to the prevalence of smartphones and other mobile devices, mHealth has the potential to provide far-reaching transformation of health care, particularly when aligned with behavior change theories utilizing trigger messages from persuasive technology [22,25-27].

mHealth apps have an advantage over computers and various print communications because they are available at nearly any time and any place (provided they are native apps) [24]. These systems can engage users (ie, patients) without requiring initiation of action by the user. However, the inclusion of theory is often overlooked in the overall design of such systems [28]. When included, behavior change theories, such as social cognitive theory (SCT; with a focus on self-efficacy) and the health belief model (HBM), are effective in user engagement of mHealth apps [19,29,30]. Persuasive technology provides a structure to allow for behavioral trigger messages and tunneling designs in such systems [9,31]. The successful integration of behavior change theories into mHealth design through the use of persuasive technology can potentially lead to reinforcement of behavior, change in attitude and belief, and ultimately a change in behavior [32]. Some of the most effective techniques include the utilization of self-monitoring components, tailoring, gamification, and utilization of push messaging for engaging patients in the management of their health care [9,11,33,34].

Messaging in Mobile Health

Persuasive technology can assist in delivering behavioral change techniques by triggering behaviors through explicit techniques such as *delivering messages at the right time to cue a specific behavior*, providing *reminders*, and using *badges as incentives for goal(s) accomplishment* [6]. These triggers can comprise of text messages, alarms, or notifications. Triggers can facilitate the performance of specific behaviors, which can provide support in accomplishing larger tasks needed in chronic disease management [5,35].

Messages can take the form of *sparks* designed for individuals who could benefit from motivational support, *facilitators* designed for those who lack ability, or *signals* designed as a simple reminder message to perform a specific behavior [35]. Although trigger messages have been used in literature, little is known about the effectiveness of specific message forms or their interactions.

Behavior Change Theory and Self-Efficacy

In life, we are challenged with individual obstacles that require us to overcome and persevere. People with chronic disease have the additional burden of self-managing their disease processes every day. To succeed in overcoming these obstacles, individuals must believe that they are capable of successfully executing certain tasks. Alfred Bandura defined self-efficacy as "the belief in one's capabilities to organize and execute the courses of action required to produce given attainments" [36]. This belief in self-efficacy is a critical component of behavior change [36,37].

Managing our health behaviors is key to reducing preventable disease and death, particularly as it relates to chronic disease [38]. The demand for those in health education and health behavior to facilitate behavior change continues to rise with a growing number of traditional and mHealth interventions to choose from [38]. This presents several problems: determining which intervention to use, which behavior change models would work best, and whether there is evidence-based medicine to support its usage. A review of the literature on preventative measures and chronic disease showcases a plethora of behavioral change models to choose from. There are a number of health behavior change models such as the HBM and SCT that focus on increasing self-efficacy to change behavior [38]. Both of these models work well in terms of helping individuals manage or control chronic diseases as they both consider self-efficacy a key concept in overall behavior change [38].

For this study, we focused on integrating a theory of behavior change (ie, SCT), the Fogg behavior model (FBM), and persuasive technology into an IHCA framework to develop an mHealth app called capABILITY [3,23,36]. We selected SCT as the researchers firmly believe that self-efficacy is critical to and has the ability to sustain behavior change through accomplishments [36,37]. The FBM was selected as it asserts that if a person was to perform a targeted behavior, he or she must have motivation, have the ability to perform a behavior, and must be triggered to perform the behavior [35]. Therefore, we developed two sets of behavioral trigger messages called sparks (designed for individuals who could benefit from motivational support) and facilitators (designed for individuals who lack ability) in an effort to enhance self-efficacy through the utilization of FBM triggers [20]. We focused on a population of individuals with type 2 diabetes as an example of a group with chronic disease that could potentially benefit from such an mHealth app. We designed capABILITY through a user-centered approach to improve self-efficacy, knowledge, and self-care in individuals with type 2 diabetes. It is important to note that only the educational content is related to type 2 diabetes, so capABILITY has the potential to be replicated in other chronic disease areas by simply changing the educational content while utilizing the same theoretical design to include the behavioral trigger messages. To that end, we will explore the following premises as a feasibility study for capABILITY: (1) explore the changes in self-efficacy, knowledge, and self-management measure scores at baseline and postintervention; (2) explore if participants would be more engaged in the use of capABILITY following a behavioral trigger; and (3) explore if participants who receive spark triggers involving motivation will engage in the utilization of capABILITY more promptly than those who receive facilitator triggers.

Methods

Study Setting

The research study was approved by the institutional review board at the University of Texas Health Science Center in Houston as well as the University of Louisiana at Lafayette. The study was conducted at a hospital system in the Gulf Coast Region. The hospital system consists of several hospitals and various ancillary health facilities (ie, physician clinics and surgical plaza).

Recruitment

Participants of the research study were either an employee or spouse of an employee from within the studied health care system. In total, 20 adult participants took part in the study. Recruitment occurred within the hospital system by emails and flyers. Participants attended a launch event where they consented, were provided education on how to use capABILITY, and had capABILITY downloaded on their device.

Focus Group Sessions: Participants and Clinical Experts

The design of mHealth apps often lacks appropriate user needs assessment [39]. According to Burke et al [40], to improve patient-centered outcomes, we must actively engage both clinicians and patients in the creation of mHealth apps that enable patients to become more effective self-managers of their chronic disease(s). With this in mind, we conducted focus groups with individuals with type 2 diabetes and clinical experts who provide their care.

The focus group (participants with type 2 diabetes and clinical experts) sessions were conducted independently. Each focus group session was conducted for 1.5 hours. The clinical expert focus group comprised 1 endocrinologist, 1 nurse practitioner, 2 registered nurses, and 3 registered dietitians. Of the 7 experts, 2 were also certified diabetes educators. In total, 9 participants with type 2 diabetes mellitus took part in the participant focus group session. These participants with type 2 diabetes were a representative sample of the population we recruited for the capABILITY study (they did not participate in the capABILITY study). The participants ranged from janitorial to clinical workers (nurses) within the hospital system.

In addition to consent documents, participants were provided with an introduction to the study that included a definition of self-efficacy. We utilized a semistructured focus group question model to stimulate open discussions based on the questions that were selected (Textbox 1) [41,42]. In addition, participants completed a demographic survey.

Textbox 1. Sample questions utilized for the focus group sessions.

1. What type of tasks do you give your patients to manage their diabetes at home? (clinical expert question)
2. What is the biggest challenge in your day-to-day diabetic self-management? (participant question)
3. What types of information should be delivered via an mHealth app? (participant and expert question)

The focus group sessions were audio-recorded and transcribed to determine common themes. Once the audio files were transcribed, we utilized qualitative software to identify key concepts and themes.

Focus Group Information

In all, two dominant themes from both groups arose, centered around critical gaps or shared beliefs. The participants identified three critical gaps in their type 2 diabetes management: health knowledge, self-management, and the financial impact of managing their disease (Multimedia Appendix 1). Shared beliefs included concepts found both within and between groups and included items such as low self-efficacy, diet struggles, and low motivation for patients with diabetes.

In addition, the focus group participants voiced a strong desire for information to be delivered in multimedia formats including short videos such as cooking tips and exercises presented visually to promote this new behavior change. They felt that this would allow them to understand better the material presented and keep them engaged in using the mHealth app.

Both the clinical expert and participant focus groups highlighted the following three areas in terms of needed education and perceived low self-efficacy: diet, exercise, and self-management. These three content areas became the core educational modules of capABILITY and were labeled as: module 1 (diet), module 2 (exercise), and module 3 (self-management).

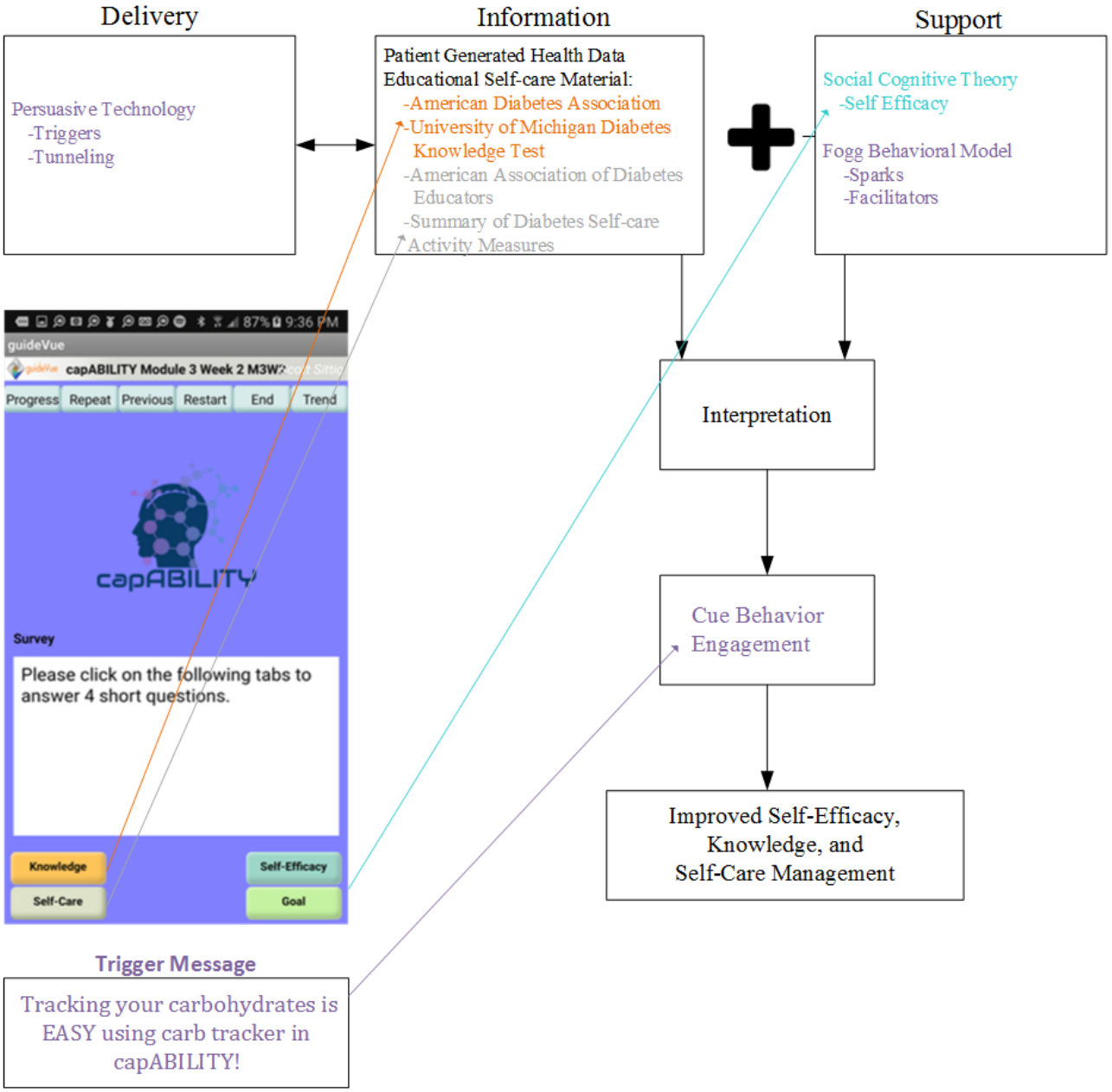
All of the experts agreed that self-efficacy plays a role in the ability of an individual with diabetes to manage their disease process. This was an important finding as our experts agreed with the published literature that improving self-efficacy is one of the keys to helping individuals manage their type 2 diabetes. In addition, 86% (6/7) of the experts stated that they have suggested a mobile device app for one of their patients (Multimedia Appendix 2).

The researchers utilized the information gained from the focus group sessions to include the critical gaps and shared beliefs to inform the user-centered design of capABILITY. This ensured that our design is reflective of the decision points we received from both key stakeholders.

capABILITY Theory Integration

The research team utilized the IHCA framework to design capABILITY. The IHCA framework allows for the delivery of health information via mHealth in combination with other theories such as behavior change or decision support [3]. Previous research has shown that IHCAs delivered through Web-based apps provide a promising way to engage users in their diabetes knowledge and self-management activities [17,19]. Building on previous IHCA frameworks and focus group sessions, we embedded patient-generated health data (PGHD) and theoretical constructs from SCT (focus on self-efficacy), FBM, and persuasive technology [3,23,36]. SCT allowed the researchers to focus on self-efficacy, which was a key decision point from the focus group sessions. Persuasive technology, FBM, and PGHD are new constructs to the IHCA framework for which we have not identified through previous works (Figure 1). The researchers feel that these are vital components to create an engaged mHealth app focused on behavioral change to improve self-efficacy, knowledge, and self-management for individuals with chronic disease (ie, type 2 diabetes). In particular, we wanted to evaluate the two types of trigger messages (sparks and facilitators) within the FBM to determine their effectiveness to deliver behavioral content within our mHealth app. We created this combination of constructs within the IHCA framework delivered through mHealth to improve self-efficacy, knowledge, and self-care management. Figure 1 depicts how the theories are integrated into the IHCA framework and ultimately into the design of capABILITY.

Figure 1. Interactive health communication application with incorporation of social cognitive theory, Fogg behavior model, and persuasive technology. IHCA: interactive health communication application; PGHD: patient-generated health data; SCT: social cognitive theory.



capABILITY Development

The next stage in the system design of capABILITY was to identify an mHealth authoring product that would allow us to incorporate our new IHCA design into the mHealth development. We ultimately decided to utilize a product called guideVUE [43]. guideVUE is an authoring app that gives you the ability to develop mHealth apps with a strong focus on knowledge transfer. guideVUE provided us the ability to embed our IHCA framework through a module (core educational content) design. We wanted to develop capABILITY with a static IHCA framework and to create three distinct educational modules focusing on diet, exercise, and self-management. This would allow our design to be replicated in other chronic disease processes by simply interchanging the educational content.

capABILITY Educational Content Development

The educational content for capABILITY was built around the three modules: module 1 (diet), module 2 (exercise), and module 3 (self-management). The development of material for each module was driven by information gathered from the focus group sessions, clinician and individual interviews, and information from the American Diabetes Association (ADA), the summary of diabetes self-care activities (SDSCA) measures, perceived diabetes self-management scale (PDSMS) and the University of Michigan Diabetes Research and Training Center's diabetes knowledge test (DKT) [11]. The majority of the educational content was retrieved from the ADA, which was transformed into media and text within capABILITY. The media files consisted of short (2-3 min) videos of one of the researchers highlighting key educational content areas such as strategies for carbohydrate counting and providing weekly content overview videos. In addition, we ensured that the videos could

be paused, rewound, and fast-forwarded so the participants could have full control of how and when they wanted to watch the videos. The text files consisted of condensed educational content from the ADA for which we also created hyperlinks in case the participants wanted to read the complete documents. This was particularly useful when we provided healthy recipes for them to utilize.

Each module within capABILITY consists of 3 weeks of unique educational material related to that particular core education module. Each week, new information is introduced in regard to that particular module. Table 1 depicts a representation of the modules and education content within capABILITY.

Table 1. capABILITY module and week classification.

Week	Module 1: Diet	Module 2: Exercise	Module 3: Self-management
1	Carbohydrate counting	Types of exercises	Diabetes facts
2	Snacks and desserts	Overcoming exercise barriers	Blood glucose
3	Diabetes superfoods	Keeping active	Medication management

The educational information gathered from the ADA was first broken down by module and, then, ultimately by week. The weekly educational topics under each specific module were created based on the information obtained from the expert and participant focus groups. To begin the classification of educational material we would use in capABILITY, we created paper folders (printed from the ADA) listed by module, then subfolders by week. This was a tedious process as we wanted to focus on the SCT construct of mastery [36,37]. Essentially, this meant that the information would be provided via capABILITY in a staggered format to promote the ideology of mastery. For instance, module 1, week 1 focused on carbohydrate counting and the ADA has a great text document discussing three strategies for better carbohydrate counting. Before transforming the paper mock-ups into the actual educational content within capABILITY, an endocrinologist and a nurse practitioner who focuses on type 2 diabetes reviewed the educational content in the folders to ensure content quality and appropriate label classification.

The development of material for each module was centered on self-efficacy, and, in particular, we utilized mastery experience, social modeling, and verbal persuasion. For example, we created knowledge questions that became increasingly more challenging as the participants gained mastery experience in a particular module, such as exercise. This technique from SCT has the strongest impact on self-efficacy belief [38]. The educational videos included statements such as *others like yourself have been successful in managing their type II diabetes*. These reinforced social modeling statements were intended to show the participants that people just like themselves have been able to manage their chronic disease successfully. Finally, we embedded verbal persuasion statements to facilitate behavior change in our trigger messages such as *Bringing HEALTHY snacks to work or on the go can help curb hunger while adding a nutritious energy boost to your day! You CAN successfully manage your diet!*

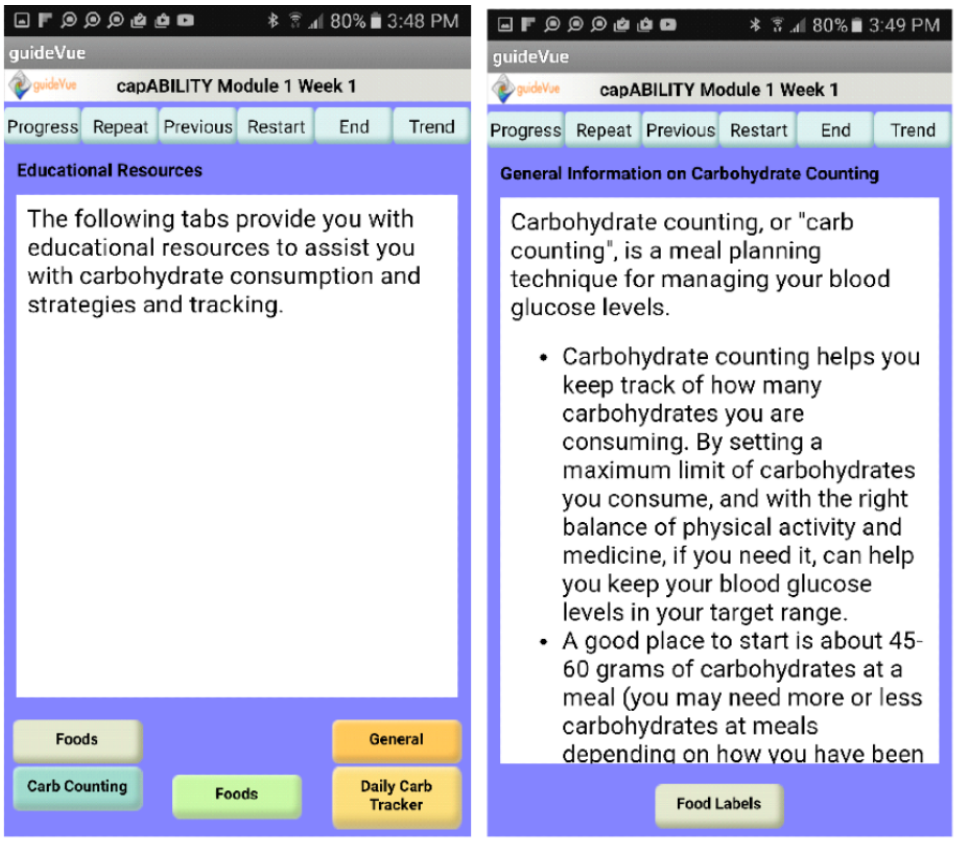
capABILITY App Development

By utilizing guideVUE, we developed and designed module 1, week 1, which would be the replicating design structure for the following 8 weeks of educational material to be delivered via capABILITY (Multimedia Appendix 3). This approach allowed us to create a flow map design infrastructure, which creates the tunneling design ensuring that each participant follows a predetermined set of screens. This was very important as previous research has shown that reducing barriers such as changes in layout is essential in trying to persuade new behaviors [44]. The only items that changed each week were the actual education content related to that week's material. This allowed the users to quickly become comfortable utilizing capABILITY and hopefully feel very comfortable utilizing the mHealth app. The premise of this design was based on the principle of *tunneling*, which is a form of persuasive technology. Through this tunneling design, we wanted to ensure that all of the users had the same experience and were exposed to specific information that they might not have seen otherwise [23]. Tunneling designs have been used to reduce cognitive load, which is important in more complex or information heavy mHealth apps such as capABILITY [23,44].

When capABILITY is first launched, the first screen the user sees is the welcome screen. This screen explains what capABILITY is and includes a capABILITY logo that the users see on most screens. At the bottom of this screen is an ID button. When the ID button is pressed, it opens a new screen for which each user can select their unique ID number from a drop-down menu. At the bottom of the ID screen is a welcome video button that leads the participant to a welcome video screen. This welcome video portrays one of the researchers as the moderator and explains what will be covered during this week's material in capABILITY. It is important to remember that only the content changes week to week so the process in which the user matriculates from screen to screen remains the same. After the user views the welcome video, he or she is able to click on the goal button at the bottom of the welcome video screen, which then leads them to a new goal's screen. At this point, the participant can then select an answer to a preformatted goal question. For example, *how many day(s) will you record your daily carbohydrate consumption*? Each week provides a new preformatted goal question for the user to answer. At the bottom of the goal screen is a resources button, which leads the user to the educational resources menu (Figure 2). This menu contains all of the educational material for the week as well as a PGHD option, which we call the tracker button. This is the main screen for which the users will spend most of their time. They are able to launch various educational, PGHD, and weekly question screens from the educational resource screen. Once the participant clicks on one of the educational resource buttons, a

new screen appears with that related content. In addition, some educational resources buttons contain multiple screens due to the educational content to be covered. Figure 2 represents screens from module 1, week 1 in capABILITY. The second screen shows what the participant would see after they click on the gray food button from the first screen.

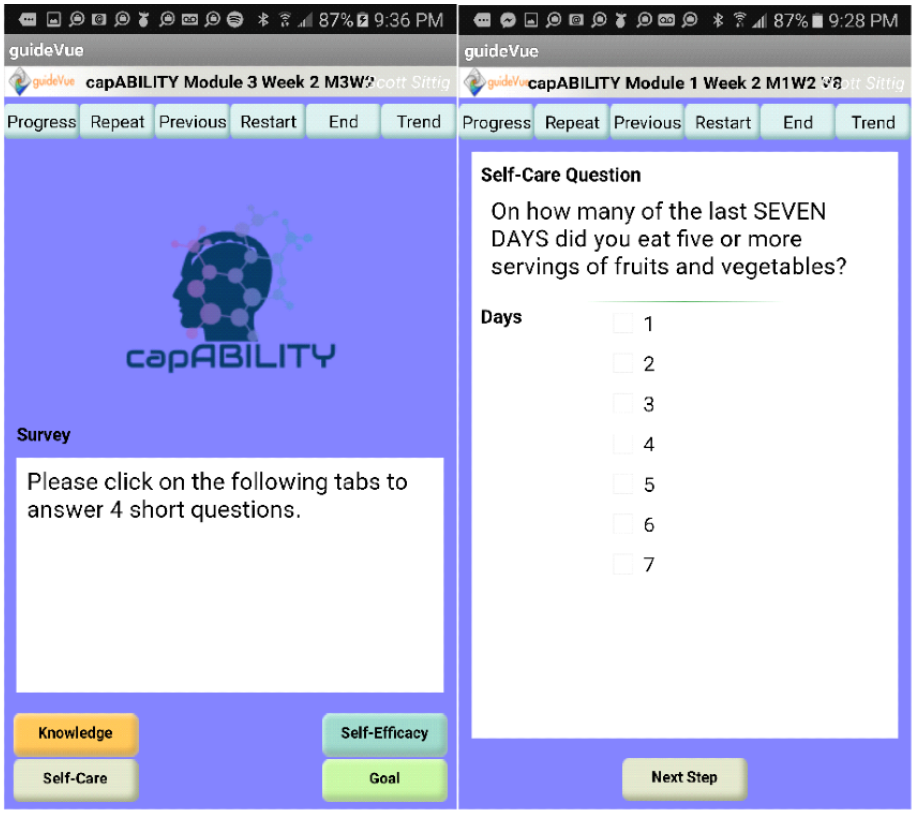
Figure 2. capABILITY resources menu.



After reviewing the educational content, the user is able to key in their PGHD by total carbohydrate consumption (module 1: Diet) by pressing the daily carb tracker button. The user is then able to select the day of the week for which they want to key in their PGHD for carbohydrate consumption. Once the user makes the day selection, a new screen appears and they are able to key in their daily carbohydrate consumption by breakfast, lunch, and dinner. The user is able to access these screens and key in PGHD at any point, which makes it easy for them to key in PGHD when it is actually being calculated.

Once the user reaches either Saturday or Sunday via the PGHD tracker, they are then prompted to open a new survey screen. At this point, a new survey screen appears for which the participant can answer four questions in total related to self-efficacy, knowledge, self-care, and goal attainment (Multimedia Appendix 4). The only question that remains constant throughout each week is the self-efficacy question *Am I generally able to accomplish my goals with respect to managing my diabetes*? The participants are able to answer the question via the following Likert Scale (ie, strongly disagree=1 to strongly disagree=5). The question is generated from the list of eight self-efficacy questions from the PDSMS [45]. The knowledge and self-care questions change each week and are related to the educational content represented that week (Figure 3). The knowledge questions are derived from the University of Michigan Diabetes Research and Training Center's DKT and are multiple-choice in nature. The self-care questions are derived from the SDSCA and are generally listed as an answer of 1 through 7 days [46]. The goal question is simply a question asking the participants if they met their goal for the week (the goals are also provided) with the following answer choices: yes, no or I'm not sure. After answering these survey questions, the participants have completed their material for the week. Each week is designed the exact same way with the exception of the PGHD content. In addition, the participants were able to key in the following PGHD components: carbohydrate consumption by meal per each day of the week, total exercise (in minutes) per day of the week, and blood glucose per each day of the week. For blood glucose PGHD, the participant can enter the blood glucose reading, pre- or postmeal and the time the blood glucose was checked.

Figure 3. capABILITY patient-generated health data tracking.



Triggers

In addition to capABILITY development, the researchers also developed spark and facilitator trigger messages to coincide with the use of capABILITY. We created three unique spark and facilitator trigger messages for each week of content within capABILITY. Essentially, we developed 27 spark triggers and facilitator triggers that would be sent to the participants (Multimedia Appendix 5 for examples). We utilized a mobile group messaging app called GroupMe, which is owned by Microsoft to deliver the trigger messages to the participants using SMS messaging. Through GroupMe, we created two mobile messaging groups called Sparks and Facilitators. This allowed us to place the participants into specific groups, which then allowed us to send either a spark or facilitator trigger message to a specific group. This design ensured that all the participants in a specific group received the exact same message and also received it at the exact same time. These messages were sent 3 days a week (eg, Tuesday, Thursday, and Saturday) at 10 AM. This time was selected to be early enough in the day to allow for an impact on the day's behavior.

capABILITY Data Capture

capABILITY was designed to capture very specific data points that would be utilized for analysis as well as user viewing (Table 2).

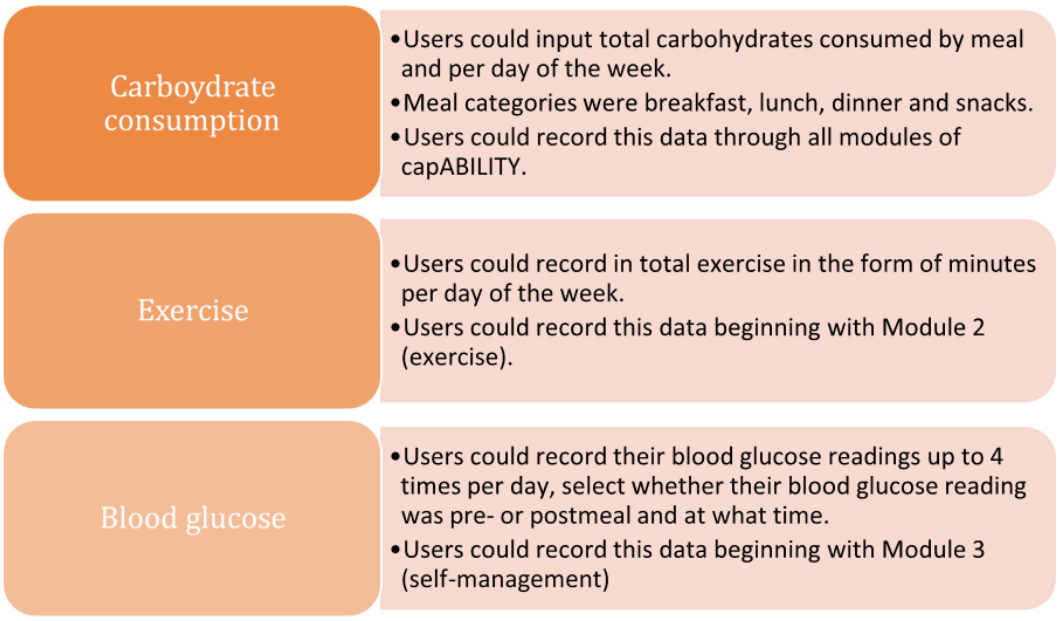
Table 2. capABILITY data capture.

Description	Data type	Collection
Participant ID	Quantitative	Each log-in
Goal statement	Quantitative and qualitative	Once per week
PGHD[a] (carbohydrates, exercise, and blood glucose)	Quantitative	Once per day
Survey questions (self-efficacy, knowledge, self-care, and goals)	Quantitative and qualitative	Once per week

[a]PGHD: patient-generated health data.

Once a user accesses a new week of material, the first screen they encounter is the goal statement screen. The goal statement changes each week and is targeted to each week's content. Goals become more challenging over the weeks as mastery develops [37]. Participants both set their goal and report whether or not they meet this milestone (answer choices were yes, no, or I'm not sure). In addition, PGHD components supported users in capturing key points such as their carbohydrate consumption, exercise, and blood glucose levels. Figure 4 represents how the participant-inputted PGHD was and how it was recorded.

Figure 4. Patient generated health data collection.



The four survey questions at the end of each week were designed to measure and evaluate behavioral and knowledge changes throughout the utilization of capABILITY. Before utilizing capABILITY, the participants completed a full self-efficacy, knowledge, and self-care measures survey. These participants would eventually complete this survey again upon completion of the capABILITY study. capABILITY was designed so that the user, at the conclusion of each week, could answer the following questions related to self-efficacy, knowledge, self-care, and goal assessment (Multimedia Appendix 4).

All of the weekly questions are derived from the list of pre- and postsurvey questions. Collecting the weekly survey data in this format was critical as it would allow us to compare a research participant's pretest and posttest data with how they were actually interacting with capABILITY weekly. These data also allow us to determine if specific types of trigger messages have an impact on self-efficacy, knowledge, and self-care.

Heuristics Evaluation

A heuristic evaluation was conducted by 2 expert reviewers who are familiar with the process. In the heuristic evaluation, the experts evaluated capABILITY for adherence to good design principles [47,48]. In total, the experts found eight violations, which were all rated to be cosmetic or minor issues. capABILITY was then redesigned to address these violations. These findings were not unexpected as the authoring tool provided templates for interaction, which had undergone extensive testing through other developments.

capABILITY User Testing

User testing was conducted with two populations: clinical users and patient users. In all, 2 clinical experts (an endocrinologist and a family nurse practitioner) and 2 individuals with type 2 diabetes (these individuals did not take part in the capABILITY intervention) were provided access to the first week of capability. Participants were asked to review the content and functionality of capABILITY. As all weeks followed the same physical structure, this limited review was believed to capture all functional issues with the system. These participants provided feedback and participated in a semistructured debriefing session. There were specific questions for the clinical experts and individuals with type 2 diabetes. Below is a sample of questions that were utilized during the interview process.

1. Do you recommend making any changes to the content? If so, what changes do you recommend?
2. Did you have any problems utilizing capABILITY or have any trouble navigating through the screens?
3. Do you recommend making any changes to capABILITY? If so, what would they be?

The 2 clinical experts felt very confident that capABILITY was providing clinically correct information about type 2 diabetes. They were both in agreement that utilizing information from the ADA as the backbone of the educational content was the best methodology. In addition, they felt strongly that allowing the user to key in PGHD data would keep them engaged and hopefully lead to them taking more responsibility in the care of their type 2 diabetes. Most of the recommendations they provided were minor or cosmetic such as the following: change the words medication adherence to medication management and add hyperlinks to critical educational resources such as carbohydrate counting strategies. We made both changes to include other cosmetic improvements as well.

The 2 individuals with type 2 diabetes felt that they were able to navigate easily through capability, and the content that was provided would help them manage their type 2 diabetes. They also stated that it was easy to key in PGHD, answer the goal question, and the weekly educational questions. Their suggested improvements included creating a button to see what content has already been viewed and to include more videos. This feedback was similarly incorporated into the design.

Utilization of capABILITY

capABILITY was a 9-week study, which covered three main diabetes content areas, which we call modules (3 weeks in each module): diet, exercise, and self-management (ie, medication adherence and glucose monitoring). Within each module, new material was delivered each week through capABILITY. Essentially, every Monday started a new week's worth of educational material that was intended to last until Sunday. In addition, a 3-crossover factor design methodology was utilized. Each participant was randomly assigned to either the control group (no triggers), spark trigger group, or facilitator trigger group. At the beginning of each module, the participants would be randomly assigned to 1 of the 3 aforementioned classification groups.

Upon conclusion of capABILITY training, consent, and completion of survey questionnaires, the participants were instructed to utilize capABILITY as they desire. capABILITY was downloaded through the app store onto each participant's phone, and then the participant utilized Wi-Fi to access the app. capABILITY was designed as weekly content files, so the participants were instructed to download each new week's worth of content each Monday. They were provided with a schedule of weekly content information for which they could refer to if needed. This process ensured that participants could not jump forward to information that was not in the canned sequence of events (referring to tunneling as a methodology of persuasive technology). Participants always had the ability to go back and view older material (weeks) and were encouraged to do so. In addition, participants were asked at the beginning of the study to complete their weekly goal, key in PGHD, and answer their weekly survey questions. This was only asked of them once at the beginning of the study as we did not want to continually remind or encourage them as this could have produced an unwarranted motivation stimulation, which would confound with the spark and facilitator trigger messages. Upon conclusion of the study, the participants completed postmeasures (self-efficacy, knowledge, and self-care) on paper. This was completed through collaboration with the nurse navigator at the hospital system.

Statistical Analysis

To determine if there was a statistically significant increase in self-efficacy, knowledge, and self-management postintervention, paired sample *t* test analyses were performed. In addition, a between-subjects one-way analysis of variance (ANOVA) was performed to determine if there was a statistically significant difference in posttest means of self-efficacy, knowledge, and self-management by the time classification in capABILITY of high, mid, and low at baseline and conclusion of the study. There were 7 participants in the high time classification that utilized capABILITY for a total of 772 min, 7 participants in the mid time classification that utilized capABILITY for a total of 299 min and 6 participants that utilized capABILITY for a total of 57 min. Paired sample *t* tests were also performed on pre- and post-: self-efficacy, knowledge, general diet, specific diet, exercise, blood glucose, and foot care. For feasibility questions 2 and 3, we followed a 3-crossover factor design and utilized a repeated measures ANOVA for analysis. The dependent variables utilized in the repeated measures ANOVA were control (C), spark trigger (S), and facilitator trigger (F). Only participants who experienced each dependent variable were utilized for the analysis (n=12; Table 3).

Table 3. Participants by trigger sequence (C=control, F=facilitator, S=spark; N=12).

Trigger sequence	Participants, n
CFS	1
CSF	4
FCS	3
FSC	1
SCF	1
SFC	2

Results

Program Outcomes

In total, 20 participants were enrolled in the study and were randomly assigned at the beginning of each module into the control, facilitator, or spark groups. Pre- and post-: self-efficacy, knowledge, and self-care measures were collected and analyzed on all 20 participants. Due to attrition during the course of the study, only 12 participants were utilized for analysis of mHealth engagement and trigger engagement. The mean age of the participants was 54.7 years (SD 10.4), and the mean number of years diagnosed with type 2 diabetes was 9 (SD 7.6). Most of the participants were female, and three-quarters of the population was white (Multimedia Appendix 6). Table 4 shows that self-efficacy, knowledge, and self-care measures all improved when posttest scores are compared with that of the pretest scores.

Table 4. Paired sample *t* test on self-efficacy, knowledge, and self-management (N=20; exploring changes in self-efficacy, knowledge, and self-management measure scores at baseline and post-capABILITY pilot study).

Outcome	Pretest, mean (SD)	Posttest mean (SD)	Change score (Δ)	2-tailed *t* test	*P* value	Cohen *d*
Self-efficacy	3.31 (0.84	3.63 (0.83)	0.32	−1.65	.12	0.38
Knowledge	0.79 (0.163)	0.82 (0.137)	0.03	−1.43	.68	0.20
General diet	3.55 (2.26)	4.37 (1.80)	0.82	−2.23	.04	0.40
Specific diet	3.13 (1.52)	3.68 (1.85)	0.55	−1.51	.15	0.32
Exercise	1.63 (1.96)	2.74 (1.75)	1.11	−3.18	.005	0.60
Blood glucose	3.39 (3.03)	4.37 (2.80)	0.98	−2.46	.02	0.36
Foot care	3.92 (2.75)	4.18 (2.29)	0.26	−0.72	.48	0.10

A paired sample *t* test was utilized on each outcome to determine the significance level pre- and post-capABILITY intervention. Results indicated statistical significance on 3 of the 7 outcomes (general diet, *P*=.04; exercise, *P*=.005; and blood glucose, *P*=.02). Table 5 displays the mean and SD of the pre- and posttest scores, including the change score (Δ) from pre-to-post and Cohen *d* effect size. If we only analyze the high and mid users (n=14) of capABILITY, we produce a statistically significant difference in self-efficacy (*P*=.008) and exercise (*P*=.01). The high users (7 in total) time range in the system was 117 to 71 min and the mid users (7 in total) time in the system ranged from 70 to 21 min. We also performed a one-way ANOVA to analyze the between-group differences (high, mid, and low) on each outcome. The one-way ANOVA did not show any statistically significant differences between groups. This could be in part to the small n within each group (high, mid, and low users).

Table 5. Paired sample *t* test on self-efficacy, knowledge, and self-management (n=14); exploring changes in self-efficacy, knowledge and self-management in only the high and mid users of capABILITY).

Outcome	Pretest, mean (SD)	Posttest, mean (SD)	Change score (Δ)	*t* test	*P* value	Cohen *d*
Self-efficacy	3.25 (0.90)	3.86 (0.75)	0.61	−3.13	.008[a]	0.74
Knowledge	0.82 (0.14)	0.85 (0.11)	0.03	−1.25	.23	0.24
General diet	3.82 (2.38)	4.96 (1.37)	1.14	−2.46	.29	0.59
Specific diet	3.14 (1.51)	3.82 (1.20)	0.68	−1.66	.12	0.50
Exercise	1.54 (2.14)	2.75 (1.86)	1.21	−2.93	.01[a]	0.60
Blood glucose	3.61 (3.25)	4.61 (2.83)	1.00	−1.88	.08	0.33
Foot care	4.22 (2.70)	4.54 (2.08)	0.32	−0.67	.51	0.13

[a]Values are statistically significant.

Engagement was operationalized by duration (ie, total time in capABILITY). To analyze duration by type of behavioral trigger (spark, facilitator, and control), the triggers were ordered in the form of a 3-factor crossover design. Figure 2 represents the ordering sequence of the participants (n=12).

A repeated measures ANOVA was run to examine the differences between the 3 different trigger types and duration. Preliminary analysis revealed that the sphericity assumption was not upheld (Mauchly's test=0.411; *P*=.01). The within-subject analysis revealed that there was not a significant effect, $F_{1,2}$=0.677; *P*=.52. In addition, descriptive statistics showed the weekly mean duration (in seconds) of time per participant in the control group (621) to be greater than spark (537) and facilitator (500) groups. Table 6 shows the engagement (duration in seconds) by module and also by trigger type. Behavioral tasks were also evaluated as participant activity within capABILITY. Behavioral tasks that participants could take part in included: setting a weekly goal, acknowledgment of meeting the goal at the end of the week, weekly PGHD input, answering a weekly self-efficacy question, answering a weekly knowledge question, and answering a weekly self-management question. Participants in the control group completed the most behavioral tasks (148), followed by particpants in the spark group (133) and finally the facilitator group (116). Participants in the spark group had the fewest incomplete behavioral tasks (44), followed by participans in the control group (50), and finally the facilitator group (51). This resulted in particpants within the spark group producing a 75.1% (133/177) behavioral task adherence which was the highest among the three groups.

Table 6. Engagement (time duration) by trigger type within each module (exploring if participants who receive spark triggers involving motivation will engage in the utilization of capABILITY more promptly than those who receive facilitator triggers).

Trigger	Duration (seconds), n (%)		
	Module 1: Diet (N=24,870)	Module 2: Exercise (N=16,201)	Module 3: Self-management (N=18,666)
Control	11,949 (48.01)	5122 (31.62)	5289 (28.33)
Facilitator	7898 (31.76)	3660 (22.59)	6475 (34.69)
Spark	5023 (20.20)	7419 (45.79)	6902 (36.98)
Total	24,870 (100.00)	16,201 (100.00)	18,666 (100.00)

Engagement was operationalized by average time from trigger delivery to capABILITY log-in. To analyze average time from trigger to capABILITY log-in by type of behavioral trigger (spark, facilitator, and control), the triggers were ordered in the form of a 3-factor crossover design (n=12).

A repeated measures ANOVA was run to examine the differences between three different triggers (spark, facilitator, and control) and the average time to log-in to capABILITY posttrigger delivery. Preliminary analysis revealed that the sphericity assumption was not upheld (Mauchly's test=0.293; P=.002). The within-subject analysis revealed that there was not a significant effect ($F_{1,2}$=0.945; P=.40). In addition, descriptive statistics showed that participants in the spark group logged in to capABILITY quicker than those in the control and facilitator groups based on the timing of trigger delivery.

As seen in the table above, the spark triggers consistently outperformed the control and facilitator triggers in terms of cueing the participants to engage with capABILITY more quickly postreceipt of a trigger. The spark trigger group produced the quickest trigger to log-in response for each module.

Participant Debriefing

A postintervention debriefing session was conducted utilizing a semistructured question format (16 questions in total). In total, 8 of the 20 participants volunteered to participate in the debriefing session, which lasted for 2 hours. The debriefing session was conducted at the main hospital in a private conference room. The main goal of the debriefing session was to find out more information on: what did the participants learn, how did capABILITY help them manage their diabetes, what aspects of capABILITY did they learn the most from, when were they most compelled to utilize capABILITY, what was their interpretation of the trigger messages, how could capABILITY be improved, and would they continue using capABILITY postintervention.

In total, 7 of the 8 participants responded to the open-ended questions, and the participants provided 97 answers for the 16 questions asked during the session.

Textbox 2 depicts a sample of questions and participant responses during the postintervention debriefing.

Textbox 2. Debriefing questions and sample answers.

What did you learn through using capABILITY?

- "I learned to identify when I was procrastinating in finding solutions to my problems."
- "How to count my carbs and the difference between good and bad carbs."

How have you changed in regard to managing your diabetes from before the study to now?

- "Increased priorities, now identifying methods to place emphasis on self-care activities."
- "More serious about diet, exercise, health in general and foot care."

What did you best learn from? Video, text, links, goals, keying of carbs, exercise or blood glucose?

- "Documenting my own information."
- "Videos."
- "Text, links, goal setting which provides a structure."

Would you like to have more control over how you receive things with regard to tailoring it to your own personal preferences (ie, set your own goals, message timing)?

- "I would have liked to see different levels as I was getting kind of bored because some of the stuff I already knew."
- "It would have been helpful to set my own goals since I have been diagnosed with type 2 diabetes for a while."

What was your overall experience with using capABILITY?

- "I lost 7 pounds during the program."
- "Although I knew most of the information presented. It made me more aware of what I was doing wrong in trying to manage my diabetes."
- "Positive and educational. Provided me with insight regarding my personal barriers to compliance."

How could capABILITY be improved?

- "The information that was inputted needs to be retrievable in an understandable format."
- "Modules for beginner, intermediate and advanced people with type II diabetes."

Would you like to continue using capABILITY?

- "I would definitely continue using it."
- "Yes, I feel this tool easily fits into my daily routines!"
- "If I did, it would help me from falling back into my old and unhealthy ways."

Discussion

Principal Findings

The results of the study show the importance of utilizing a user-centered design approach to incorporate behavioral theoretical constructs into a framework that integrates the needs of the end-users and clinical experts. Data analysis showed improvements in self-efficacy, knowledge, and self-management; however, not all of them showed a statistically significant change from pre-to-post intervention. When the data were analyzed with all participants (N=20), only the survey measures of general diet, specific diet, and blood glucose showed statistically significant improvements. Essentially, the utilization of capABILITY produced the most significant changes in self-management. When the data from only the high and mid users (n=14) of capABILITY was analyzed, a statistically significant difference in self-efficacy and general diet (survey data) was observed. The significance of self-efficacy changed considerably from the first analysis (N=20) of P=.12 to the second analysis (n=14) of P=.008. Therefore, the data hint that there is a difference between groups and that the more time spent utilizing capABILITY, the more appreciable improvement in self-efficacy may be expected. Although there were improvements in knowledge outcome scores, these gains did not produce a statistically significant difference from preintervention to postintervention. This was not surprising as we learned through our earlier focus group sessions and postintervention debriefing session that knowledge was not directly correlated to self-efficacy. Some participants who scored very high on their knowledge tests also scored very low on their self-efficacy survey. These participants told us that although they have a high knowledge level, they did not feel they could add something else to their already full load of being a provider, spouse, or parent. These participants were typically those with a clinical education background such as nursing. In addition, the knowledge scores overall were high to start, so there was not much room for growth. Finally, we determined that there was not a statistically significant difference in postmeasure

(survey data) outcomes between the three time classification groups (high, mid, and low).

The parameters for understanding engagement and behavioral trigger messages were that a participant must be active in capABILITY and receive all three types of triggers (control, spark, and facilitator). This parameter reduced our sample size to 12 due to attrition throughout the course of the study. In addition, we operationalized engagement as the duration of time spent utilizing capABILITY. We also used descriptive data from behavioral tasks within capABILITY such as setting a weekly goal, acknowledgment of meeting the goal at the end of the week, weekly PGHD input, answering a weekly self-efficacy question, answering a weekly knowledge question, and answering a weekly self-management question.

The repeated measures ANOVA showed that there was not a significant within-subject effect between the trigger types and duration. The results also showed that when participants were in the control group they engaged (duration) with capABILITY more than when they were in the spark or facilitator trigger group. Overall, participants in the control group utilized capABILITY for 22,360 seconds, as compared with 18,033 seconds for participants in the facilitator group and 19,344 for participants in the spark. Every 3 weeks (start of new module), the participants were randomized into 1 of the 3 trigger groupings. At the start of the study (module 1), there were 5 participants in the control group, 4 in the facilitator group, and 3 in the spark group. As we ended up with 12 participants for this analysis, the start of the randomized grouping order may have impacted engagement as a whole. As seen in Table 6, duration time in module 1 far exceeded duration time in modules 2 and 3. This is common at the beginning of a study; however, there were 5 participants in the control to start the study, compared with only three in the spark group. In modules 2 and 3, the participants in the spark group outperformed (more duration time in capABILITY) those in the control and facilitator groups. It is plausible that, if the randomized trigger groupings started out with the same number of participants in the spark group as the control group, we would see the spark group with the largest overall duration time. Although it would not be statistically significant, it would be an important descriptive data finding.

In addition to the engagement (duration) analysis, a descriptive analysis was conducted on behavioral tasks within capABILITY. The control group completed the most behavioral tasks (148), followed by the spark group (133), then the facilitator group (116). As stated above, this could be linked to more participants starting module 1 in the control group. Although the control group completed the most behavioral tasks, the spark group had the highest adherence percentage to completing the behavioral tasks.

The repeated measures ANOVA showed that there was not a significant within-subject effect between the trigger types and the duration of time between trigger delivery to participant log-in of capABILITY. Although the results were not statistically significant, the spark triggers did produce the fastest response from trigger to capABILITY log-in.

The fact that the spark triggers engaged the participants to log in to capABILITY at a much quicker response rate is a very important finding. The FBM states that for a person to accomplish a specific behavioral task, the following must occur: be motivated, have the ability or capacity to perform the behavior, and to be triggered to perform the behavior [35]. Spark triggers could be the missing link in the attempt to cue individuals to perform a specific behavior within a given amount of time.

It is interesting to note that both the spark and facilitator triggers outperformed the control group in engaging the participants to log in to capABILITY quicker; however, individuals in the control group actually spent more time using capABILITY. We feel this confirms that the triggers (in particular, the spark) cue an individual to accomplish a task, but do not necessarily improve their engagement as time spent within a system. This is evidenced in a study by Weymann et al [17], where a tailored IHCA designed for individuals with chronic diseases showed that the participants spent significantly more time in the system compared with the control group; however, it did not lead to more knowledge or patient empowerment. Combining a tailored IHCA mHealth app with spark triggers could potentially improve both engagement in the system as well as behavioral outcomes. Future work with larger sample sizes should explore this idea further to determine if spark developed triggers engage users to cue a particular behavior quicker. In addition, motivation scales should be used to ascertain initial baseline scores to determine the effect this has on triggers, especially spark triggers (these have a focus on motivation).

Limitations

The primary limitation of this study was the small sample size, which did not produce a large enough statistical power for us to detect statistically significant changes in the engagement of behavioral triggers. Second, all the participants in the study were employed full time with benefits, which may not fully represent a typical chronic disease population. Third, individual differences in motivation and extrinsic factors, such as the timing of the study, may have an impact. Studies of larger samples and longer designs would address these concerns. Finally, participant attrition did impact the ability to conduct robust statistical analysis.

Conclusions

We utilized a user-centered design process, which incorporated individuals with type 2 diabetes and clinical experts. This process is critical in understanding the decision points of the key stakeholders to integrate into an mHealth design. The IHCA framework allows for the inclusion of behavior change theories and persuasive technology (including trigger messages) to be integrated into an mHealth system design for individuals with chronic disease. Our work suggests that self-efficacy, knowledge, and self-management may be improved through utilization of a theory-driven mHealth app. Future work should focus on replicating this model in other chronic diseases with larger sample sizes to determine if self-efficacy, knowledge, and self-management can be improved.

In addition, our work implies that spark triggers have the ability to cue specific individual actions quicker than facilitator triggers or simply no triggers at all. This is an important discovery in the area of consumer informatics as we may be able to design triggers through a targeted population-based approach instead of individualized tailored triggers. The creation of population-based spark triggers for chronic disease could be an effective approach to cueing positive behavioral tasks for large populations at a time through mHealth. This could become a powerful tool that could be utilized in accountable care organizations, managed care organizations, large health care systems, or population health management at any level. It is to be noted that our research findings in the area of spark triggers differs from the idea in the FBM that individuals may be more tolerant of facilitators or reminders over the course of time [35]. The 9-week study showed that spark triggers continually cued participants to engage with capABILITY at the beginning and conclusion of the study. From these findings, we feel that trigger messages, which contain motivation (sparks) in the form of pleasure, hope, and social acceptance, cue actions quicker than facilitator messages or simple reminders.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Focus group theme classifications.
[DOCX File , 14 KB - mhealth_v8i3e15927_app1.docx]

Multimedia Appendix 2
Expert and participant social demographic survey.
[DOCX File , 15 KB - mhealth_v8i3e15927_app2.docx]

Multimedia Appendix 3
capABILITY module 1, week 1 flowchart.
[DOCX File , 77 KB - mhealth_v8i3e15927_app3.docx]

Multimedia Appendix 4
Module 1 weekly survey questions.
[DOCX File , 13 KB - mhealth_v8i3e15927_app4.docx]

Multimedia Appendix 5
Spark and facilitator trigger messages: module 1, weeks 1 and 2.
[DOCX File , 14 KB - mhealth_v8i3e15927_app5.docx]

Multimedia Appendix 6
Participant demographic and pre- and posttest data.
[DOCX File , 19 KB - mhealth_v8i3e15927_app6.docx]

Multimedia Appendix 7
CONSORT‐EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 4408 KB - mhealth_v8i3e15927_app7.pdf]

References

1. Smith A. Pew Research Center. 2017 Jan 12. Record Shares of Americans Now Own Smartphones, Have Home Broadband URL: http://www.pewresearch.org/fact-tank/2017/01/12/evolution-of-technology/ [accessed 2018-03-12]
2. Jepson RG, Harris FM, Platt S, Tannahill C. The effectiveness of interventions to change six health behaviours: a review of reviews. BMC Public Health 2010 Sep 8;10:538 [FREE Full text] [doi: 10.1186/1471-2458-10-538] [Medline: 20825660]
3. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. Cochrane Database Syst Rev 2005 Oct 19(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
4. Kitsiou S, Paré G, Jaana M, Gerber B. Effectiveness of mHealth interventions for patients with diabetes: an overview of systematic reviews. PLoS One 2017;12(3):e0173160 [FREE Full text] [doi: 10.1371/journal.pone.0173160] [Medline: 28249025]

5. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014 Jun 5;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
6. Chatterjee S, Price A. Healthy living with persuasive technologies: framework, issues, and challenges. J Am Med Inform Assoc 2009;16(2):171-178 [FREE Full text] [doi: 10.1197/jamia.M2859] [Medline: 19074300]
7. Ahola AJ, Groop P. Barriers to self-management of diabetes. Diabet Med 2013 Apr;30(4):413-420. [doi: 10.1111/dme.12105] [Medline: 23278342]
8. Gao J, Wang J, Zhu Y, Yu J. Validation of an information-motivation-behavioral skills model of self-care among Chinese adults with type 2 diabetes. BMC Public Health 2013 Feb 4;13:100 [FREE Full text] [doi: 10.1186/1471-2458-13-100] [Medline: 23379324]
9. Goyal S, Morita P, Lewis GF, Yu C, Seto E, Cafazzo JA. The systematic design of a behavioural mobile health application for the self-management of type 2 diabetes. Can J Diabetes 2016 Feb;40(1):95-104. [doi: 10.1016/j.jcjd.2015.06.007] [Medline: 26455762]
10. Arora S, Peters AL, Agy C, Menchine M. A mobile health intervention for inner city patients with poorly controlled diabetes: proof-of-concept of the TExT-MED program. Diabetes Technol Ther 2012 Jun;14(6):492-496. [doi: 10.1089/dia.2011.0252] [Medline: 22524591]
11. Mayberry LS, Mulvaney SA, Johnson KB, Osborn CY. The MEssaging for diabetes intervention reduced barriers to medication adherence among low-income, diverse adults with type 2. J Diabetes Sci Technol 2017 Jan;11(1):92-99 [FREE Full text] [doi: 10.1177/1932296816668374] [Medline: 27595710]
12. Fortmann AL, Gallo LC, Garcia MI, Taleb M, Euyoque JA, Clark T, et al. Dulce Digital: an mHealth SMS-based intervention improves glycemic control in Hispanics with type 2 diabetes. Diabetes Care 2017 Oct;40(10):1349-1355 [FREE Full text] [doi: 10.2337/dc17-0230] [Medline: 28600309]
13. Hors-Fraile S, Schneider F, Fernandez-Luque L, Luna-Perejon F, Civit A, Spachos D, et al. Tailoring motivational health messages for smoking cessation using an mHealth recommender system integrated with an electronic health record: a study protocol. BMC Public Health 2018 Jun 5;18(1):698 [FREE Full text] [doi: 10.1186/s12889-018-5612-5] [Medline: 29871595]
14. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
15. Alahäivälä T, Oinas-Kukkonen H. Understanding persuasion contexts in health gamification: A systematic analysis of gamified health behavior change support systems literature. Int J Med Inform 2016 Dec;96:62-70. [doi: 10.1016/j.ijmedinf.2016.02.006] [Medline: 26944611]
16. Korinek EV, Phatak SS, Martin CA, Freigoun MT, Rivera DE, Adams MA, et al. Adaptive step goals and rewards: a longitudinal growth model of daily steps for a smartphone-based walking intervention. J Behav Med 2018 Feb;41(1):74-86. [doi: 10.1007/s10865-017-9878-3] [Medline: 28918547]
17. Weymann N, Dirmaier J, von Wolff A, Kriston L, Härter M. Effectiveness of a web-based tailored interactive health communication application for patients with type 2 diabetes or chronic low back pain: randomized controlled trial. J Med Internet Res 2015 Mar 3;17(3):e53 [FREE Full text] [doi: 10.2196/jmir.3904] [Medline: 25736340]
18. Weymann N, Härter M, Petrak F, Dirmaier J. Health information, behavior change, and decision support for patients with type 2 diabetes: development of a tailored, preference-sensitive health communication application. Patient Prefer Adherence 2013;7:1091-1099 [FREE Full text] [doi: 10.2147/PPA.S46924] [Medline: 24174871]
19. Weymann N, Härter M, Dirmaier J. A tailored, interactive health communication application for patients with type 2 diabetes: study protocol of a randomised controlled trial. BMC Med Inform Decis Mak 2013 Feb 13;13:24 [FREE Full text] [doi: 10.1186/1472-6947-13-24] [Medline: 23406466]
20. Research 2 Guidance. 2017. 325,000 Mobile Health Apps Available in 2017 – Android Now the Leading mHealth Platform URL: https://research2guidance.com/325000-mobile-health-apps-available-in-2017/ [accessed 2019-02-11]
21. Andrew A, Borriello G, Fogarty J. Toward a systematic understanding of suggestion tactics in persuasive technologies. In: de Kort Y, IJsselsteijn W, Midden C, Eggen B, Fogg BJ, editors. Persuasive Technology. Berlin, Heidelberg: Springer; 2007:259-270.
22. Connelly KH, Faber AM, Rogers Y, Siek KA, Toscos T. Mobile applications that empower people to monitor their personal health. Elektrotech Inftech 2006;123(4):124-128. [doi: 10.1007/s00502-006-0326]
23. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco, CA: Morgan Kaufmann; 2003.
24. Revere D, Dunbar PJ. Review of computer-generated outpatient health behavior interventions: clinical encounters 'in absentia'. J Am Med Inform Assoc 2001;8(1):62-79 [FREE Full text] [doi: 10.1136/jamia.2001.0080062] [Medline: 11141513]
25. Henkemans OA, van der Boog PJ, Lindenberg J, van der Mast CA, Neerincx MA, Zwetsloot-Schonk BJ. An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. Technol Health Care 2009;17(3):253-267. [doi: 10.3233/THC-2009-0545] [Medline: 19641261]

26. Giunti G. 3MD for Chronic Conditions, a Model for Motivational mHealth Design: Embedded Case Study. JMIR Serious Games 2018 Aug 24;6(3):e11631 [FREE Full text] [doi: 10.2196/11631] [Medline: 30143476]
27. Orji R, Moffatt K. Persuasive technology for health and wellness: State-of-the-art and emerging trends. Health Informatics J 2018 Mar;24(1):66-91. [doi: 10.1177/1460458216650979] [Medline: 27245673]
28. van Vugt M, de Wit M, Cleijne WH, Snoek FJ. Use of behavioral change techniques in web-based self-management programs for type 2 diabetes patients: systematic review. J Med Internet Res 2013 Dec 13;15(12):e279 [FREE Full text] [doi: 10.2196/jmir.2800] [Medline: 24334230]
29. McMahon S, Vankipuram M, Hekler EB, Fleury J. Design and evaluation of theory-informed technology to augment a wellness motivation intervention. Transl Behav Med 2014 Mar;4(1):95-107 [FREE Full text] [doi: 10.1007/s13142-013-0221-4] [Medline: 24653780]
30. Tufano JT, Karras BT. Mobile eHealth interventions for obesity: a timely opportunity to leverage convergence trends. J Med Internet Res 2005 Dec 20;7(5):e58 [FREE Full text] [doi: 10.2196/jmir.7.5.e58] [Medline: 16403722]
31. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: A scoping review. J Med Internet Res 2017 Aug 1;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
32. Oinas-Kukkonen H, Harjumaa M. Towards Deeper Understanding of Persuasion in Software and Information Systems. In: Proceedings of the First International Conference on Advances in Computer-Human Interaction. 2008 Presented at: ACHI'08; February 10-15, 2008; Sainte Luce, Martinique. [doi: 10.1109/achi.2008.31]
33. Lee H, Ghebre R, Le C, Jang YJ, Sharratt M, Yee D. Mobile phone multilevel and multimedia messaging intervention for breast cancer screening: pilot randomized controlled trial. JMIR Mhealth Uhealth 2017 Nov 7;5(11):e154 [FREE Full text] [doi: 10.2196/mhealth.7091] [Medline: 29113961]
34. Freyne J, Yin J, Brindal E, Hendrie GA, Berkovsky S, Noakes M. Push notifications in diet apps: influencing engagement times and tasks. Int J Hum–Comput Interact 2017;33(10). [doi: 10.1080/10447318.2017.1278896]
35. Fogg BJ. A Behavior Model for Persuasive Design. In: Proceedings of the 4th International Conference on Persuasive Technology. USA: ACM; 2009 Presented at: Persuasive'09; April 26 - 29, 2009; Claremont, CA p. 1-7. [doi: 10.1145/1541948.1541999]
36. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52:1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
37. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
38. Glanz K, Rimer BK, Viswanath K. Health Behavior and Health Education: Theory, Research, and Practice. Fourth Edition. San Francisco, CA: Jossey-Bass; 2008.
39. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
40. Burke SD, Sherr D, Lipman RD. Partnering with diabetes educators to improve patient outcomes. Diabetes Metab Syndr Obes 2014;7:45-53 [FREE Full text] [doi: 10.2147/DMSO.S40036] [Medline: 24550679]
41. Wang J, Matthews JT. Chronic disease self-management: views among older adults of Chinese descent. Geriatr Nurs 2010;31(2):86-94. [doi: 10.1016/j.gerinurse.2009.11.001] [Medline: 20381709]
42. Wang J, Siminerio LM. Educators' insights in using chronicle diabetes: a data management system for diabetes education. Diabetes Educ 2013;39(2):248-254. [doi: 10.1177/0145721713475844] [Medline: 23427239]
43. Iyengar M, Florez-Arango J, Garcia C. GuideView: A System for Developing Structured, Multimodal, Multi-platform Persuasive Applications. In: Proceedings of the 4th International Conference on Persuasive Technology. USA: ACM; 2009 Presented at: Persuasive'09; April 26 - 29, 2009; California, Claremont, USA p. 1-5. [doi: 10.1145/1541948.1541990]
44. Heffernan KJ, Chang S, Maclean ST, Callegari ET, Garland SM, Reavley NJ, et al. Guidelines and recommendations for developing interactive eHealth apps for complex messaging in health promotion. JMIR Mhealth Uhealth 2016 Feb 9;4(1):e14 [FREE Full text] [doi: 10.2196/mhealth.4423] [Medline: 26860623]
45. Wallston KA, Rothman RL, Cherrington A. Psychometric properties of the Perceived Diabetes Self-Management Scale (PDSMS). J Behav Med 2007 Oct;30(5):395-401. [doi: 10.1007/s10865-007-9110-y] [Medline: 17522972]
46. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care 2000 Jul;23(7):943-950 [FREE Full text] [doi: 10.2337/diacare.23.7.943] [Medline: 10895844]
47. Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. J Biomed Inform 2003;36(1-2):23-30. [doi: 10.1016/s1532-0464(03)00060-1] [Medline: 14552844]
48. Nielsen J, Molich R. Heuristic Evaluation of User Interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. USA: ACM; 1990 Presented at: CHI'90; April 1 - 5, 1990; Washington, Seattle, USA p. 249-256. [doi: 10.1145/97243.97281]

Abbreviations

ADA: American Diabetes Association
ANOVA: analysis of variance
DKT: diabetes knowledge test
FBM: Fogg behavior model
HBM: health belief model
IHCA: interactive health communication application
mHealth: mobile health
PDSMS: perceived diabetes self-management scale
PGHD: patient-generated health data
SCT: social cognitive theory
SDSCA: summary of diabetes self-care activities measure

Edited by G Eysenbach; submitted 19.08.19; peer-reviewed by D Kaufman, V Mylonopoulou, J Floch; comments to author 18.09.19; revised version received 13.11.19; accepted 10.02.20; published 16.03.20.

Please cite as:
Sittig S, Wang J, Iyengar S, Myneni S, Franklin A
Incorporating Behavioral Trigger Messages Into a Mobile Health App for Chronic Disease Management: Randomized Clinical Feasibility Trial in Diabetes
JMIR Mhealth Uhealth 2020;8(3):e15927
URL: http://mhealth.jmir.org/2020/3/e15927/
doi:10.2196/15927
PMID:32175908

Original Paper

Pilot Study of a Multilevel Mobile Health App for Substance Use, Sexual Risk Behaviors, and Testing for Sexually Transmitted Infections and HIV Among Youth: Randomized Controlled Trial

David Cordova[1], PhD; Jaime Munoz-Velazquez[1], MS; Frania Mendoza Lua[2], MSW; Kathryn Fessler[3], MD, PhD; Sydni Warner[1]; Jorge Delva[4], PhD; Nicole Adelman[3], MPH; Youth Leadership Council[5]; Angela Fernandez[1], PhD, MPH, LCSW; Jose Bauermeister[6], PhD, MPH

[1]School of Social Work, University of Michigan, Ann Arbor, MI, United States

[2]School of Social Service Administration, University of Chicago, Chicago, IL, United States

[3]Corner Health Center, Ypsilanti, MI, United States

[4]School of Social Work, Boston University, Boston, MA, United States

[5]See Acknowledgments

[6]School of Nursing, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
David Cordova, PhD
School of Social Work
University of Michigan
1080 S University Ave
Room 2772
Ann Arbor, MI, 48109
United States
Phone: 1 734 763 6201
Email: cordovad@umich.edu

Abstract

Background: Preventing and reducing substance use disorders, sexually transmitted infections (STIs)/HIV, and teen pregnancy, and the associated risk behaviors (ie, substance use and sexual risk behaviors) among youth remain public health priorities in the United States. Equally important is improving the uptake of STI/HIV testing among the youth. Mobile health (mHealth) apps may be a solution to ameliorate these public health concerns; however, few mHealth preventive interventions have demonstrated efficacy in reducing substance use or sexual risk behaviors or improving the uptake of STI/HIV testing among the youth, particularly in clinic settings.

Objective: This small-scale study aimed to examine the feasibility of conducting a pilot randomized controlled trial (RCT). We evaluated the effects of Storytelling 4 Empowerment (S4E), relative to enhanced usual practice, on the potential mechanisms by which behavior change occurs, namely clinician-youth risk communication, prevention knowledge, and substance use and sexual risk refusal self-efficacy. We also assessed the ability to measure targeted outcomes of past 30-day substance use (ie, alcohol, tobacco, and other drug use), condomless sex, and alcohol or drug use before sex, as well as the uptake of STI/HIV testing.

Methods: Employing community-based participatory research principles, 50 youths aged 13 to 21 years were recruited from a youth-centered community health clinic in Southeast Michigan, randomized sequentially to either S4E or enhanced usual practice, and assessed at baseline, immediately postintervention, and 30 days postintervention. S4E consists of 3 modules, including alcohol and drug use, tobacco, and STI/HIV.

Results: Relative to youth in the enhanced usual practice group, S4E participants demonstrated higher youth-clinician risk communication (mean 3.22, SD 1.67) and increases in prevention knowledge (Δ score mean 0.36, SD 0.51) and self-efficacy (Δ score mean 0.16, SD 0.47). In addition, youth in the S4E group showed reductions in the proportions of past 30-day overall substance use (Cohen *h*=0.71, 95% CI 0.15 to 1.27), as well as past 30-day alcohol (Cohen *h*=0.71, 95% CI 0.15 to 1.27), tobacco (Cohen *h*=0.17, 95% CI −0.39 to 0.73), and drug use (Cohen *h*=1.28, 95% CI 0.72 to 1.84). The results also suggest a reduction in the proportion of youths who reported past 30-day condomless sex (Cohen *h*=0.18, 95% CI −0.38 to 0.74) and alcohol use

before sex (Cohen *h*=0.44, 95% CI −0.12 to 1.00). Finally, the findings also demonstrated an increase in the proportion of youths who reported STI/HIV testing over time (Cohen *h*=0.16, 95% CI −0.39 to 0.72).

Conclusions: The findings suggest the feasibility of a small-scale pilot RCT. S4E demonstrated shifts in the hypothesized direction, reducing substance use, sexual risk behaviors, and improving the uptake of STI/HIV testing among youth in a clinic setting. The findings suggest that a larger RCT may be warranted.

Trial Registration: ClinicalTrails.gov NCT03855410, https://clinicaltrials.gov/ct2/show/NCT03855410.

(JMIR Mhealth Uhealth 2020;8(3):e16251) doi:10.2196/16251

KEYWORDS

youth; mHealth; illicit drugs; sex behavior; HIV; primary care

Introduction

The Prevalence of Sexually Transmitted Infections/HIV and Associated Risk Behaviors Among the Youth

Sexually transmitted infections (STIs), including HIV, and teen pregnancy, remain significant public health concerns among youth in the United States [1-4]. STIs and HIV infection have been linked to infertility, cancer, and increasing vulnerability to opportunistic infections [5,6]. In addition, teen pregnancy has been linked to low income, poverty, and low educational attainment [7]. Therefore, preventing and reducing STIs, HIV, and teen pregnancy, as well as associated risk behaviors such as substance use and sexual risk behaviors, remain critical public health priorities.

Substance use and sexual risk behaviors may directly or indirectly increase the risk of STIs, HIV, and teen pregnancy. These behaviors often increase during adolescence, highlighting the need to intervene in these behaviors throughout this developmental period of increased vulnerability [8-10]. National epidemiologic data suggest that substance use behaviors, including alcohol, tobacco, and other drug use, are widespread among the youth [11]. Substance use behaviors often parallel other risk behaviors [12] and have been linked to increased sexual risk behaviors among the youth [8,13,14]. In addition to preventing and reducing substance use and sexual risk behaviors, STI and HIV testing are key strategies to reduce the high rates of STI/HIV infection among the youth [15,16]. Despite the Centers for Disease Control and Prevention guidelines, many youths are not routinely screened for STIs [17], and 90.7% of 9th- to 12th-grade students report having never been tested for HIV [11]. Therefore, there remains an urgent need to identify settings and tools that may be leveraged to improve the uptake of STI/HIV testing among youth.

Leveraging Youth-Centered Community Health Clinics and Mobile Health Apps to Prevent Sexually Transmitted Infections/HIV Among the Youth

Youth-centered community health clinics may be an ideal setting to engage youth in prevention services. Evidence supports that compared with adult-focused clinics and AIDS service organizations, youth are more likely to seek substance use and sexual risk prevention and risk reduction services from youth-centered community health clinics [18,19]. However, relatively few interventions have been developed and tested in clinic settings [20,21]. Leveraging clinic settings, in combination with technology, may have great utility in identifying substance use, sexual risk behaviors, and STI/HIV testing solutions for youth.

Mobile health (mHealth) refers to medical or public health initiatives and practices supported by mobile devices such as tablets and the internet [22]. Among a limited yet growing body of research, mHealth interventions have been pilot-tested and demonstrated positive shifts in reducing substance use, sexual risk behaviors, or increasing STI/HIV testing among youth [23-28]. For example, researchers have shown that brief mHealth interventions reduce marijuana use among youths aged 15 to 24 years at 3 months postintervention [24] and heavy alcohol consumption among young bisexual men at 3 months postintervention [26]. Other research has shown that brief mHealth interventions can improve the uptake of HIV [23,25,28,29] and STI testing [23] and decrease the frequency of condomless sex [25].

Limitations of Scientific Knowledge on Mobile Health Preventive Interventions

Although scientific advancements on mHealth preventive interventions have been made, several important limitations exist. First, few mHealth preventive interventions focus on substance use and concurrent risk behaviors (ie, sexual risk behaviors) in younger adolescents (aged <18 years) [25,30], missing the opportunity to affect a key developmental period of enhanced risk-taking [31]. Second, interventions targeting sexual risk behaviors and uptake of STI/HIV testing have focused primarily on young men who have sex with men [23,25,29], with few interventions focused on other vulnerable populations. Indeed, stark HIV disparities among men who have sex with men exist, accounting for 87% of new HIV diagnoses among youths aged 13 to 24 years [32]. Also important are racial and ethnic minority youth and adolescent women who constitute additional vulnerable populations of youth [32-34]. Third, to date, interventions have focused primarily on linking the youth to STI/HIV testing sites [23,29], with relatively few preventive interventions focused on the youth once they arrive at the clinic. Although drawing youth to the clinic is an important first step, it does little good if effective prevention services are not provided once the youth arrive at the clinic. Simply focusing on drawing the youth to clinics may create missed opportunities for engaging the youth in additional prevention strategies, particularly among those who are unaware of their engagement in risky behaviors [30]. To address these limitations, we conducted a small-scale randomized controlled

trial (RCT) to pilot-test the feasibility of a multilevel mHealth preventive intervention among a diverse sample of youth in a clinic setting.

The Storytelling 4 Empowerment Mobile Health Preventive Intervention App

Employing community-based participatory research (CBPR) principles [35], we developed Storytelling 4 Empowerment (S4E) [30,36]. Guided by ecodevelopmental [37] and empowerment theories [38], S4E aims to reduce substance use, sexual risk behaviors, and improve uptake of STI/HIV testing through improving clinician-youth risk communication, prevention knowledge, and self-efficacy [30]. This is accomplished through a multilevel mHealth app that provides interactive, targeted, and tailored content focused on the prevention of substance use and sexual risk behaviors. This mHealth app is then followed up with a clinician-initiated prevention and risk reduction face-to-face encounter, providing clinicians the opportunity to reinforce content provided to youth during their interaction with the mHealth app. Because S4E has been shown to have a positive user experience, an effective user interface, and high feasibility and acceptability among both youth and clinicians [30,39], a next important step is to conduct a small-scale pilot RCT to determine the feasibility of S4E and examine shifts in potential mechanism of change and the ability to measure substance use, sexual risk behaviors, and uptake of STI/HIV testing among a diverse sample of youth. We believe our S4E multilevel approach may offer advantages over other approaches for several reasons. First, our intervention was developed with and for the targeted community. For example, youth helped steer the development of S4E with regard to the user experience and user interface [36]. Researchers affirm that community-engaged research employing CBPR principles may lead to enhanced uptake of, and optimally efficacious, preventive interventions [35]. Second, our intervention is developmentally and culturally congruent, utilizing spaces and tools that align with youth perspectives. Specifically, youth-centered clinics are safe spaces for many youths, thereby providing a potentially high-impact context to improve public health. Furthermore, approximately 95% of the youth report having access to mobile devices [40], which may be leveraged to deliver risk behavior solutions to this vulnerable population. Finally, our S4E approach was grounded in decades of science and informed by prevention principles [41]. For example, S4E is theory driven, targets multiple levels, and focuses on multiple risk behaviors that often co-occur [30,36,39].

Purpose of the Study

The purpose of this study was to conduct a small-scale pilot RCT to determine the feasibility of S4E, relative to enhanced usual practice. We evaluated changes in the potential mechanisms of change, namely clinician-youth risk communication, prevention knowledge, and self-efficacy over time. We also assessed the ability to measure substance use, sexual risk behaviors, and uptake of STI/HIV testing over time among a diverse sample of youth in a clinic setting. Given the small-scale pilot nature of our study and sample size, statistical significance was deemphasized. Rather, our goal was to assess feasibility and establish the critical parameters necessary to inform a larger future RCT.

Methods

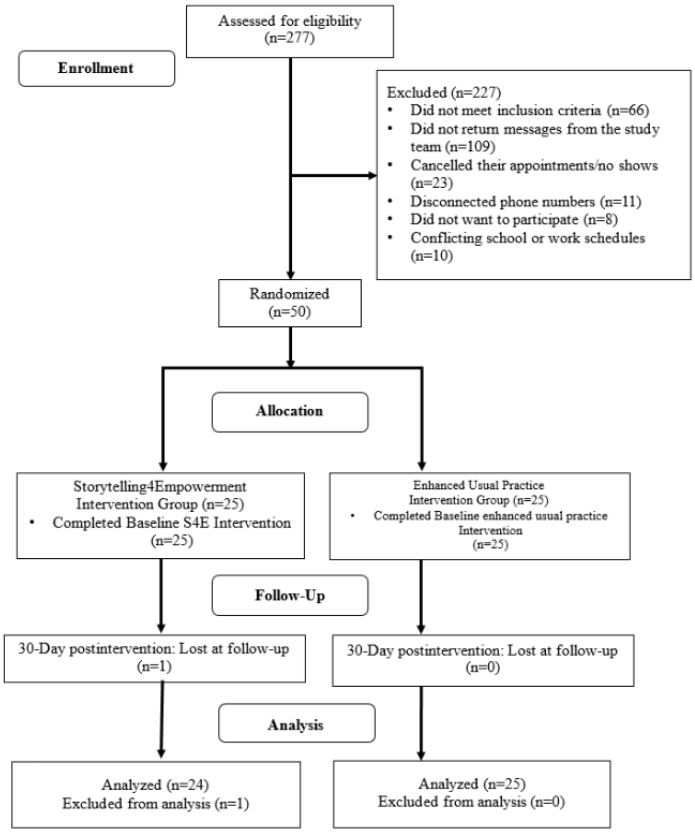
Participants

We recruited youth and clinicians between October 2016 and July 2017 from a youth-centered community health clinic located in Southeast Michigan that offers a full range of health care, mental health, and supportive services to young people as they transition to adulthood.

Youth recruitment occurred during the clinic's health appointment reminder phone calls. To be eligible for this study, youth had to (1) be aged between 13 and 21 years, (2) live in Southeast Michigan without plans to move out of the area during the study period, (3) have a scheduled appointment with a participating clinician, and (4) report no prior history of psychiatric hospitalization. Of the 277 youths who were screened for eligibility, 211 met the inclusion criteria. Of these, 109 youths did not return calls or messages left by the study team, 23 canceled or did not show for their scheduled health appointment, 11 had disconnected or incorrect phone numbers, 10 reported conflicting schedules with school or work, and 8 refused to participate. Therefore, of the 211 eligible youths, 50 were successfully recruited (see Figure 1).

Clinician recruitment occurred during weekly clinic staff meetings, and all clinicians at the health clinic were eligible to participate if they (1) worked in Southeast Michigan and (2) worked with our target population. The study team approached 8 clinicians, of which 7 agreed to participate. Both youths and clinicians who expressed interest in participating in the study were contacted by the study team to screen for eligibility, to enroll, and to complete study consent protocols. To protect the confidentiality of the youth aged between 13 and 17 years, a waiver of parental permission was obtained. This waiver was in accordance with the state of Michigan's Title X Program and Public Health Code, MCL 333.6121, which states that a minor aged 17 years or younger can consent to sexual and reproductive health and substance use services without parental knowledge. Both participating youths older than 18 years and clinicians were presented with consent through a comprehensive written waiver of documentation that did not require a signature from the participant or legally authorized representative while containing all the elements of informed consent required by the Health and Human Services' regulations and policy.

Figure 1. Consolidated Standards of Reporting Trials table.



Sample Characteristics

In line with other research pilot testing mHealth preventive interventions [24,26], the sample size for this study was 50, with 25 youth participants randomized to the S4E group and 25 randomized to the enhanced usual practice group. Of the 50 youths, 41 (82%) identified as female, followed by 4 (8%) males, and 4 (8%) transmales, and 1 refused to respond. The mean age of the youths was 18.82 years (SD 2.1, range 13-21). The racial composition of these 50 youths consisted of 23 (46%) non-Hispanic white, 21 (42%) black, 1 (2%) Native American, 4 (8%) ascribing to more than one race, and 1 (2%) selecting Other. Regarding the youths' educational attainment, 36% (18/50) of youths reported having completed some college, whereas 30% (15/50) reported having completed high school. The remaining 34% (17/50) of youths reported having completed a grade between 7th and 11th.

Among the 7 clinicians who agreed to participate, 6 (86%) identified as female, with a mean age of 43.14 years (SD 7.95, range 34-56), and 5 (71%) were non-Hispanic white, followed by 1 (14%) Hispanic or Latino and 1 (14%) Asian. They reported an average of 10.86 years (SD 7.45, range 1-22) of medical practice in their respective specialties: 71% (5/7) practiced family medicine and 29% (2/7) were pediatricians. Finally, 71% (5/7) of the clinicians reported having lived in the area where they work for more than 10 years, and 29% (2/7) reported having lived in the area for fewer than 10 years.

Study Design

This study employed community-based participatory research principles [35]. A youth leadership council, clinic director, and staff were involved in all aspects of this research, including preparing and submitting the proposal to fund this study, identifying the target population, developing the study design, and disseminating the study findings (eg, publications). This study consisted of a 2 (group) × 2 (time) small-scale pilot RCT. Youth participants were randomly assigned to either the S4E experimental group or enhanced usual practice control group via sequential randomization [42]. Data were collected on tablets and captured using Research Electronic Data Capture, a Health Insurance Portability and Accountability Act–compliant, web-based app that is hosted on secure servers at the University of Michigan Medical School. To reduce potential bias, eligible participants completed health surveys that included questions regarding substance use, sexual risk behaviors, and STI/HIV testing practices, before randomization. All youths arrived 1 hour before their scheduled health care appointment to have the study explained to them in detail, to provide consent, and to complete baseline assessments in a reserved room, all of which took approximately 30 min. Youths participated in the intervention (S4E or enhanced usual practice) in a reserved room with internet connection for approximately 30 min, while they waited for their health appointment. Participants completed the S4E intervention on tablets provided to them by the research team. The S4E mHealth version tested in this study was developed for Apple's operating system (iOS). Because this was a phase I/II pilot study, we had participants complete the S4E intervention in the clinic to have a more controlled environment. Youth participants were assessed at baseline before their health appointment, immediately postintervention, and 30 days postbaseline. Clinicians were assessed at baseline and immediately postintervention for each health appointment.

We retained 49 participants at our 30-day follow-up (49/50, 98% retention rate). Youth participants received a total of US $60 in incentives, corresponding to US $20 at baseline and US $40 at the 30-day follow-up. Through a collaborative process, the clinic and research team decided to provide a US $2000 incentive to benefit the entire community health clinic for providing us with meeting space and for the clinicians' time on the project, rather than give individual incentives to the participating clinicians.

Study Groups

Storytelling 4 Empowerment Group

The S4E intervention content was generated through community-university research involving youth-led groups in conjunction with scientific prevention principles [30]. Theory-driven, S4E takes a multilevel approach and is guided by an ecodevelopmental [37] and empowerment framework [38]. Youths in the S4E group received targeted, tailored prevention content based on their responses to the S4E risk behavior assessment, which includes the Car, Relax, Alone, Forget, Friends, Trouble screener [43]. This assessment is intended to identify the youths' specific risk behaviors based on the past year and lifetime reports of substance use, sexual risk behaviors, and past 6-month STI/HIV testing practices (Multimedia Appendix 1). From an empowerment perspective, the scores prompt risk-specific interactive prevention content (eg, short animated storytelling scenarios, interactive diagram of body health activities) for the S4E youth (Figure 2). This interactive content is delivered via 3 modules (ie, alcohol and drug use, tobacco, and STI/HIV) and aims to improve prevention knowledge, self-efficacy, and refusal skills while linking youth to important adult figures. From an ecodevelopmental perspective, the clinician-facing app contained the participants' health appointment information, their assigned group, and their S4E risk assessment responses. Both the youth and clinician S4E apps work synergistically, providing clinicians access to the youths' risk responses, motivational interviewing scripts to facilitate clinician-youth communication, and resources to local services that are based on the youths' risk behaviors (Figure 3). Overall, the S4E intervention aims to improve substance use and sexual risk prevention knowledge, self-efficacy, and refusal skills, as well as to facilitate clinician-youth risk communication. By focusing on these malleable factors, the overarching goal of the S4E intervention is disease prevention (ie, substance use disorders, STI/HIV) and health promotion (ie, STI/HIV testing) through reductions in substance use, sexual risk behaviors, and improvements in STI/HIV testing among youth.

Figure 2. Storytelling 4 Empowerment animated storytelling scenario.

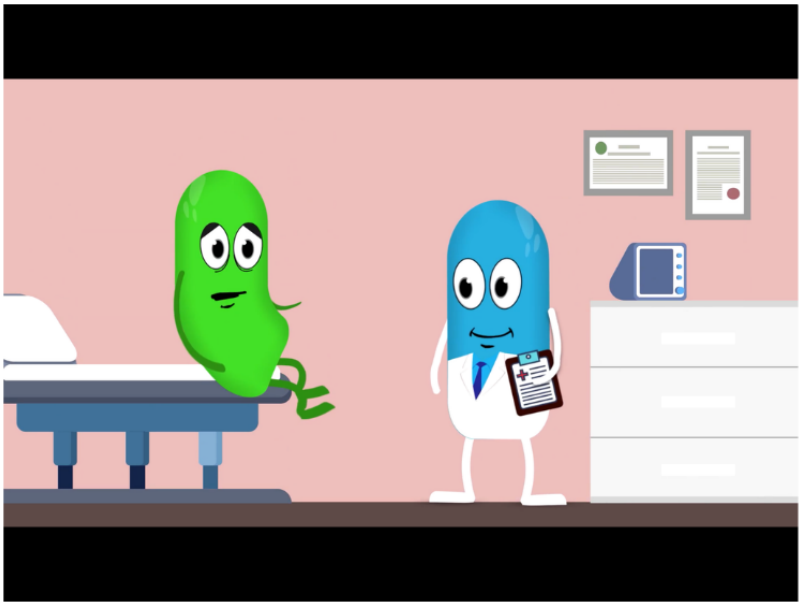


Figure 3. Clinician-facing Storytelling 4 Empowerment application highlighting youth risk behaviors.

Enhanced Usual Practice Group

The participating youth-centered community health clinic's usual practice consists of primary care, mental health, sexual and reproductive health, substance use prevention, support, and education (eg, Women, Infants, and Children Supplemental Food Program), and gender-affirming health care to youths aged 12 to 25 years. Participants in both the S4E group and enhanced usual practice received the clinic's usual services. In addition, participants in the enhanced usual practice group received a binder with a printed PDF version of the S4E tobacco module content to view before their health care appointment. The materials consisted of epidemiologic statistics and prevention health information related to tobacco and cigarette use. The Youth Leadership Council strongly recommended an enhanced usual practice control group so that all youth participants in the trial would receive some form of additional prevention services.

Measures

Demographics

Youth and clinicians completed a demographic survey that asked their age, ethnicity, race, and gender identity. In addition, youth reported educational attainment, and clinicians reported their medical specialty, years of clinical experience, and length of time residing in the area where they practiced.

Clinician-Youth Communication (Immediately Postintervention)

Both clinicians (Cronbach alpha=.81, 9-item) and youth (Cronbach alpha=.95, 9-item) interactions during the clinic visit were assessed via items extracted and adapted from the matched pair instrument (MPI) [44] immediately postintervention. These items assessed the process and content of the communication, including the language used and behaviors performed by clinicians related to substance use and sexual risk behaviors services. A sample statement for clinicians is as follows: *Encouraged the patient to express his or her thoughts concerning drug use and/or sexual risk behaviors.* A sample statement for youths is as follows: *My doctor encouraged me to express my thoughts concerning drug use and/or sexual risk behaviors.* Both clinicians and youth responded to items in the measure using a 6-point Likert-type scale (0=*not applicable*, 5=*strongly agree*). To make the MPI sum scores more interpretable, the scores were rescaled to their original 6-point metric by dividing the sum total by the total number of items before significance testing.

Self-Efficacy Outcomes

Substance Use Refusal Skills

Youths' substance use refusal skills were assessed through 2 separate items on a 4-point scale (1=*very hard* to 4=*not very hard*). Sample questions included the following: *Pretend your best friend offered you a drink of beer or wine and you did not want it. How hard would it be to say no?* and *Pretend your best friend offered you some marijuana and you did not want it. How hard would it be to say no?*

Sexual Risk Behavior Refusal Skills

One item was used to assess youths' sexual risk behavior refusal skills. The statement read the following: *I can say no to sex if my partner and I do not have a condom.* Responses were on a 5-point agreement scale (1=*strongly disagree*, 5=*strongly agree*).

Prevention Knowledge Outcomes

Substance Use Prevention Knowledge

Youths' knowledge about alcohol or drugs and tobacco products was assessed through 2 separate items. Items included the following: *If I use alcohol or drugs, I will have more health problems than other people* and *If I use tobacco products, I will have more health problems.* Responses were on a 4-point agreement scale (1=*disagree a lot* to 4=*agree a lot*).

Sexual Risk Prevention Knowledge

Sexual risk prevention knowledge was measured through 2 separate items related to the effectiveness of condom use. Items included the following: *Condoms help prevent pregnancy,* and *If I have sex without a condom, I am likely to get HIV/STIs.*

Responses were on a 5-point agreement scale (1=*strongly disagree*, to 5=*strongly agree*).

Behavioral Outcomes

Substance use behaviors, sexual risk behaviors, and HIV/STI testing were dichotomized for analysis, whereby *0* was used for *No* (ie, no risk present), and *1* indicated *Yes* (ie, risk was present) for the item in question.

Substance Use Behaviors (Baseline and 30-Day Postbaseline)

Youths' lifetime and past 30-day substance use behaviors were assessed using items adapted from the Monitoring the Future study [45,46]. Sample items included the following: *Have you ever had any beer, wine, wine cooler, or liquor to drink*? and *Have you had more than a few sips of alcohol on more than one occasion during the past 30 days?*

Sexual Risk Behaviors (Baseline and 30-Day Postbaseline)

Participants' lifetime and past 30-day sexual risk behaviors were assessed using items extracted from the Sexual Behavior Instrument [47]. Sample items include *Have you ever had vaginal, anal, or oral sex without using a condom?* and *In the past 30 days, about how often have you had vaginal, anal, or oral sex without using a condom?*

Sexually Transmitted Infection and HIV Testing (Baseline, Immediately Postintervention, 30-Day Postbaseline)

We assessed youths' lifetime and most recent STI and HIV testing. Example questions included the following: *Have you ever been tested for HIV?* and *Did you receive an STI test?*

Analytic Strategy

Given the modest sample size and goals of a pilot RCT, significance testing by group was deemphasized [23]. Rather, we determined whether outcomes shifted in the hypothesized direction and gathered the necessary parameters to use in a larger RCT in the future [48]. The data analytic strategy consisted of 4 steps. First, we conducted a descriptive statistical analysis on demographic and outcome variables at baseline and used chi-square tests and analysis of variance to test for significant group differences at baseline. Second, we conducted chi-square tests and 2-tailed *t* tests to determine if there were significant differences by group in attrition and elapsed time between baseline and 30-day follow-up assessments. Third, we determined group differences in change of potential mechanisms of change (ie, self-efficacy, prevention knowledge, clinician-youth risk communication) over time. Finally, we assessed the ability to measure between-group differences (ie, S4E vs control) in the change of reported substance use, sexual risk behaviors, and STI/HIV testing over time. We also assessed the ability to measure within-group differences in outcomes among participants in the S4E intervention group. Differences were determined using proportion change (Δ %) for categorical outcomes and mean change (Δ score, absolute net gains) over time for continuous outcomes. We report the observed effect sizes for the outcome change difference scores by group using Cohen *d* for continuous outcomes and Cohen *h* for binary outcomes, which measure the difference between 2 proportions (h=2arcsin × (sqrt P_1) − 2arcsin × (sqrt P_2), where P_1=proportion 1 and P_2=proportion 2). Effect sizes were estimated as small (d/h<.20), medium (.20≥d/h≤.45), and large (d/h>.45) to observe the magnitude of differences [49]. All analyses were performed in SPSS version 24 [50], with the exception of Cohen *h* power calculations, which were performed in R's version 3.5.2 PWR package [51].

Results

Comparability of Groups

As shown in Table 1, chi-square tests and analyses of variance results suggest no significant S4E vs control group differences at baseline on any demographic characteristic (eg, race), lifetime or past 30-day substance use, sexual risk behaviors, and lifetime STI/HIV testing. The absence of significant differences in these variables at baseline suggests that our trial's randomization procedures were successful. The median number of days between baseline and follow-up was 31 (mean 32.63, SD 8.62); no significant differences by group in the number of days between baseline and 30-day follow-up assessments were observed using a 2-tailed *t* test, $t_{46.36}$=0.42, P=.68. In addition, no significant differences in attrition by group were observed, χ^2_1=0.0, P=.92.

Table 1. Baseline comparisons by group on demographic and behavioral outcomes (N=50).

Outcomes	S4E[a] (n=25)	Control (n=25)	*t* test/chi-square (*df*)	*P* value
Age (years), mean (SD)	18.6 (2.15)	19.0 (2.19)	−0.58 (47.98)[b]	.56
Gender, n (%)				
Female	21 (84)	20 (80)	1.0 (2)	.61
Male	2 (8)	2 (8)	1.0 (2)	.61
Transmale	1 (4)	3 (12)	1.0 (2)	.61
Transfemale	—[c]	—	1.0 (2)	.61
Refuse to answer	1 (4)	—	1.0 (2)	.61
Race, n (%)				
Black	11 (44)	10 (40)	2.4 (4)	.66
White	10 (40)	13 (52)	2.4 (4)	.66
Native American	1 (4)	—	2.4 (4)	.66
More than one race	2 (8)	2 (8)	2.4 (4)	.66
Other	1 (4)	—	2.4 (4)	.66
Lifetime substance use, n (%)				
Lifetime alcohol use	21 (84)	21 (84)	—	—
Lifetime tobacco use	16 (64)	17 (68)	0.9 (1)	.77
Lifetime other drug use	14 (56)	13 (52)	0.1 (1)	.78
Past 30-day substance use, n (%)				
Past 30-day substance use (ATOD[d])	15 (60)	16 (64)	0.1 (1)	.77
Past 30-day alcohol use	10 (40)	14 (56)	1.3 (1)	.26
Past 30-day tobacco use	7 (28)	11 (44)	1.4 (1)	.23
Past 30-day other drug use	12 (48)	7 (28)	1.3 (1)	.26
Lifetime sexual risk behaviors, n (%)				
Lifetime condomless sex	20 (80)	20 (80)	—	—
Lifetime alcohol use before sex	9 (36)	14 (56)	2.0 (1)	.16
Life drug use before sex	7 (28)	8 (32)	0.2 (1)	.69
Past 30-day sexual risk behaviors, n (%)				
Past 30-day condomless sex[e]	13 (62)	13 (62)	0.2 (1)	.67
Past 30-day alcohol use before sex[e]	2 (10)	3 (14)	0.2 (1)	.67
Past 30-day drug use before sex[e]	3 (14)	4 (19)	0.1 (1)	.73
Lifetime HIV/STI[f] testing	22 (88)	20 (80)	0.6 (1)	.44

[a]S4E: Storytelling 4 Empowerment.

[b]*t* test for age; rest are chi-square values.

[c]Data are not applicable.

[d]ATOD: alcohol, tobacco, other drug use.

[e]Among sexually active youth.

[f]STI: sexually transmitted infection.

Intervention Effects on Clinician-Youth Risk Communication

Immediately postintervention, youths in the S4E group reported higher levels of clinician-youth risk communication (mean 3.22, SD 1.67), relative to the youths in the control group (mean 2.96, SD 1.63; $t_{46.77}$=0.56; *P*=.58). Although these group differences were not statistically significant, the estimated effect size (Cohen *d*=0.16, 95% CI −0.41 to 0.72) yielded a small effect size. Similarly, S4E clinicians reported higher levels of

clinician-youth risk communication (mean 3.47, SD 1.13), relative to clinicians in the control group immediately postintervention (mean 3.23, SD 1.02; $t_{45.98}$=0.75; *P*=.45). Although these differences were not statistically significant, the estimated effect size (Cohen *d*=0.22, 95% CI −0.34 to 0.79) yielded a small value.

Intervention Effects on Self-Efficacy

As shown in Table 2, youths in the S4E group reported greater change scores in substance use self-efficacy alcohol refusal (Δ score mean 0.22, SD 0.67), relative to the youths in the control group (Δ score mean 0.16, SD 0.55; $t_{42.79}$=0.32; *P*=.75; Cohen *d*=0.10) and for drug refusal (Δ score mean 0.09, SD 0.68), relative to the youths in the control group (Δ score mean 0.08, SD 0.70; $t_{44.51}$=0.05; *P*=.96; Cohen *d*=0.01). Yet, both the S4E group (Δ score mean 0.08, SD 0.78) and the control group (Δ score mean 0.13, SD 0.92) showed an increase in sex self-efficacy ($t_{44.06}$=−0.19; *P*=.85; Cohen *d*=−0.06). In contrast to the between-group effects, intervention effects within the S4E group (Table 3) for both substance use items, alcohol refusal (Δ score mean 0.21; t_{22}=1.55, *P*=.14; Cohen *d*=0.38) and drug use refusal (Δ score mean 0.09; t_{21}=0.62; *P*=.54; Cohen *d*=0.12), yielded small to medium effect sizes. Similarly, sexual risk self-efficacy change scores within the S4E group (Δ score mean 0.08; t_{23}=0.53; *P*=.60; Cohen *d*=0.10) yielded small effect sizes.

Intervention Effects on Prevention Knowledge

As shown in Table 2, youths in the S4E group reported higher overall gains in substance use prevention knowledge for tobacco use (Δ score mean 0.30, SD 0.77), relative to the youths in the control group (Δ score mean 0.16, SD 0.80; $t_{45.93}$=0.64; *P*=.53; Cohen *d*=0.18). The S4E group reported similar gains of prevention knowledge for alcohol or drug use (Δ score mean 0.35, SD 0.65), relative to the youths in the control group (Δ score mean 0.20, SD 1.00; $t_{41.46}$=0.61; *P*=.54; Cohen *d*=0.18). In addition, S4E youths reported overall gains for sexual risk prevention knowledge (pregnancy prevention, Δ score mean 0.08, SD 0.65; STI/HIV prevention, Δ score mean 0.25, SD 1.22), relative to the control group (pregnancy prevention, Δ score mean 0.00, SD 0.41; $t_{38.29}$=0.53; *P*=.60; Cohen *d*=0.15 and STI/HIV prevention, Δ score mean −0.17, SD 1.74; $t_{43.35}$=0.96; *P*=.34; Cohen *d*=0.28). Both outcomes yielded small to medium effect sizes. In contrast to the between-group effects, intervention effects within the S4E group (Table 3) for both substance use prevention knowledge (tobacco use Δ score mean 0.34; t_{22}=2.58; *P*=.02; Cohen *d*=0.49 and alcohol or drug use Δ score mean 0.31; t_{22}=1.91; *P*=.07; Cohen *d*=0.50) and the sexual risk prevention knowledge-pregnancy prevention (Δ score mean 0.08; t_{23}=0.62; *P*=.54; Cohen *d*=0.10) and STI/HIV prevention (Δ score mean 0.25; t_{23}=1.00; *P*=.32; Cohen *d*=0.21) yielded small to medium effect sizes.

Table 2. Self-efficacy and prevention knowledge by group.

Outcomes	S4E[a], Δ score, mean (SD)	Control, Δ score, mean (SD)	Independent *t* test (*df*)	Cohen's *d* (S4E vs control)	Cohen *d* 95% CI
Substance use self-efficacy					
Alcohol refusal	0.22 (0.67)	0.16 (0.55)	0.32 (42.79)	0.10	−0.47 to 0.66
Drug refusal	0.09 (0.68)	0.08 (0.70)	0.05 (44.51)	0.01	−0.56 to 0.59
Sexual risk self-efficacy					
Condomless sex	0.08 (0.78)	0.13 (0.92)	−0.19 (44.06)	−0.06	−0.61 to 0.50
Substance use prevention knowledge					
Use of tobacco products	0.30 (0.77)	0.16 (0.80)	0.64 (45.93)	0.18	−0.40 to 0.76
Use of alcohol or drugs	0.35 (0.65)	0.20 (1.00)	0.61 (41.46)	0.18	−0.39 to 0.74
Sexual risk prevention knowledge					
Pregnancy prevention	0.08 (0.65)	0.00 (0.41)	0.53 (38.29)	0.15	−0.41 to 0.71
STI[b]/HIV prevention	0.25 (1.22)	−0.17 (1.74)	0.96 (41.35)	0.28	−0.28 to 0.84

[a]S4E: Storytelling 4 Empowerment.

[b]STI: sexually transmitted infection.

Table 3. Storytelling 4 Empowerment intervention effects on self-efficacy and prevention knowledge (n=25).

Outcomes	S4E[a] baseline, mean (SD)	S4E follow-up, mean, (SD)	Δ score mean	Paired *t* test (*df*)	Cohen *d*	Cohen *d* 95% CI
Substance use self-efficacy						
Alcohol refusal	3.57 (0.59)	3.78 (0.52)	0.21	1.55 (22)	0.38	−0.94 to 0.19
Drug refusal	3.50 (0.86)	3.59 (0.67)	0.09	0.62 (21)	0.12	−0.68 to 0.44
Sexual risk self-efficacy						
Condomless sex	3.25 (0.85)	3.33 (0.76)	0.08	0.53 (23)	0.10	−0.66 to 0.46
Substance use prevention knowledge						
Use of tobacco products	3.09 (0.79)	3.43 (0.59)	0.34	2.58 (22)	0.49	−0.10 to 1.07
Use of alcohol or drugs	3.30 (0.56)	3.61 (0.58)	0.31	1.91 (22)	0.50	−0.07 to 1.07
Sexual risk prevention knowledge						
Pregnancy prevention	3.50 (0.72)	3.58 (0.50)	0.08	0.62 (23)	0.10	−0.43 to 0.69
STI[b]/HIV prevention	2.67 (1.24)	2.92 (1.10)	0.25	1.00 (23)	0.21	−0.35 to 0.77

[a]S4E: Storytelling 4 Empowerment.

[b]STI: sexually transmitted infection.

Between-Group Intervention Effects on Substance Use Behaviors

Overall, participant reports of substance use at baseline were not significantly different (Table 1). However, the S4E group reported a greater reduction in any substance use (ie, ATOD) relative to the control group (3/25, 12% vs 0/25, 0%; Table 4). Although chi-square testing (χ^2_2=4.5; *P*=.10) suggests that these proportion differences between groups were marginally significant, the estimated proportion change effect size difference between groups (Cohen *h*=0.71) yielded a large effect size. We then deconstructed past 30-day substance use into past 30-day alcohol, tobacco, and other drug use to determine between-group intervention effects on each of these outcomes.

Table 4. Past 30-day behavior outcome proportion change scores by group.

Outcomes	Δ S4E[a]	Δ Control	Cohen *h* (S4E vs control)	Cohen *h*, 95% CI
Substance use, n (%)				
Substance use (ATOD[b])	3 (−12)	0 (0)	0.71	0.15 to 1.27
Alcohol use	3 (−12)	0 (0)	0.71	0.15 to 1.27
Tobacco use	2 (−8)	1 (−4)	0.17	−0.39 to 0.73
Other drug use	3 (−12)	2 (+8)	1.28	0.72 to 1.84
Sexual risk behaviors, n (%)[c]				
Condomless sex	2 (−10)	1 (−5)	0.18	−0.38 to 0.74
Alcohol use before sex	1 (− 5)	0 (0)	0.44	−0.12 to 1.00
Drug use during sex	2 (−10)	2 (−10)	N/A[d]	N/A

[a]S4E: Storytelling 4 Empowerment.

[b]ATOD: alcohol, tobacco, other drug use.

[c]Sexual risk behaviors are based on responses from sexually active participants (n=42).

[d]Not applicable.

Alcohol Use

Overall, 84% (42/50) of participants reported lifetime alcohol use. Relative to participants in the control group, S4E group participants reported a greater reduction in past 30-day alcohol use at 30-day follow-up (3/25, 12% vs. 0/25, 0%). Although chi-square testing (χ^2_2=3.9; *P*=.14) suggests that this proportion change difference was not statistically significant between groups, the estimated proportion change effect size between groups (Cohen *h*=0.71) yielded a large effect size (Table 4).

Tobacco Use

Overall, 66% (33/50) of participants reported lifetime tobacco use. Participants in the S4E group reported a greater decrease in tobacco use, as compared with participants in the control group at 30-day follow-up (2/25, 8% vs 1/25, 4%). Although

chi-square testing (χ^2_1=0.4; *P*=.50), suggests that these proportion differences were not statistically significant, the estimated effect size difference across groups (Cohen *h*=0.17) yielded a small effect size (Table 4).

Other Drug Use

Overall, 54% (27/50) of participants reported lifetime drug use. Participants in the S4E group reported a reduction (3/25, 12%) in drug use at 30-day follow-up, relative to an increase in drug use (2/25, 8%) among participants in the control group. Although chi-square testing (χ^2_2=2.9; *P*=.23) suggests that these proportion differences were not statistically significant, the estimated effect size difference across groups (Cohen *h* =1.28) yielded a large effect size (Table 4).

Between- Group Intervention Effects on Sexual Risk Behaviors

Condomless Sex

Overall, 80% (40/50) of participants reported engaging in lifetime condomless sex. Relative to sexually active participants in the control group (n=21), sexually active S4E group participants (n=21) reported a greater reduction in past 30-day condomless sex at 30-day follow-up (2/21, 10% vs 1/21, 5%). Although chi-square testing (χ^2_2=0.2; *P*=.91) suggests that these proportion differences were not statistically significant, the estimated effect size difference between groups (Cohen *h*=0.18) yielded a small effect size (Table 4).

Alcohol Use Before Sex

Overall, 46% (23/50) of participants reported lifetime alcohol use before sex. Relative to sexually active participants in the control group (n=21) who reported no change, sexually active S4E group (n=21) participants reported a reduction in alcohol use before sex at 30-day follow-up (1/21, 5% vs 0/21, 0%). Although chi-square testing (χ^2_2=2.3; *P*=.32) suggests that these proportion differences were not statistically significant, the estimated effect size difference between groups (Cohen *h*=0.44) yielded a medium effect size (Table 4).

Drug Use Before Sex

Overall, 15 (30%) participants reported lifetime drug use before sex. Both the control and S4E group participants reported similar reductions in drug use before sex at 30-day follow-up (2/21, 10% vs 2/21, 10%; Table 4).

Within-Group Intervention Effects on Substance Use Behaviors

As shown in Table 5, of the 25 youths in the S4E group, 12% (3/25) reported a decrease of any substance use from baseline assessment to 30-day follow-up (χ^2_1=10.9; *P*<.001; Cohen *h*=0.24). We then separated past 30-day substance use into past 30-day alcohol, tobacco, and other drug use to determine within-group intervention effects on each of these outcomes. S4E intervention effects for alcohol use from baseline assessment to 30-day follow-up showed a 12% (3/25) decrease in alcohol use (χ^2_1=3.6; *P*=.06; Cohen *h*=0.24). Similarly, 8% (2/25) of the S4E participants reported a decrease in tobacco use from baseline assessment to 30-day follow-up (χ^2_1=14.6; *P*<.001; Cohen *h*=0.19). In addition, 12% (3/25) of S4E youth reported a decrease in other drug use from baseline assessment to 30-day follow-up (χ^2_1=17.0; *P*<.001; Cohen *h*=0.24). Although chi-square significance was deemphasized, S4E intervention effects on substance use behaviors yielded small to medium effect sizes.

Table 5. Past 30-day Storytelling 4 Empowerment intervention effects on behaviors from baseline to follow-up (n=25).

Outcomes	S4E[a] baseline, n (%)	S4E follow-up, n (%)	Δ score, n (%)	Cohen *h*	Cohen *h*, 95% CI
Substance use behavior					
Substance use (ATOD[b])	15 (60)	12 (48)	3 (−12)	0.24	−0.32 to 0.80
Alcohol use	10 (40)	7 (28)	3 (−12)	0.24	−0.31 to 0.81
Tobacco use	7 (28)	5 (20)	2 (−8)	0.19	−0.37 to 0.74
Other drug use	12 (48)	9 (36)	3 (−12)	0.24	−0.32 to 0.80
Sexual risk behaviors[c]					
Condomless sex	13 (62)	11 (52)	2 (−10)	0.19	−0.37 to 0.75
Alcohol use before sex	2 (10)	1 (5)	1 (−5)	0.18	−0.38 to 0.74
Drug use before sex	3 (14)	1 (5)	2 (−10)	0.33	−0.23 to 0.89

[a]S4E: Storytelling 4 Empowerment.

[b]ATOD: alcohol, tobacco, other drug use.

[c]Sexual risk behaviors are based on responses from sexually active participants (n=21).

Within-Group Intervention Effects on Sexual Risk Behaviors

As shown in Table 5, within-group intervention effects on S4E sexual risk behaviors from baseline assessment to 30-day follow-up had small to medium effect sizes that helped establish differences. Specifically, S4E reports of condomless sex decreased by 10% (2/21) from baseline assessment to 30-day follow-up (χ^2_1=2.9; *P*=.09; Cohen *h* =0.19). Moreover, decreases in alcohol use before sex (1/25, 5%; χ^2_1=8.97; *P*=.003; Cohen

h=0.24) and drug use before sex (2/25, 10%; χ^2_1=8.47; *P*=.004; Cohen *h* =0.33) were observed.

Intervention Effects on Sexually Transmitted Infections/HIV Testing

We sought to determine whether STI/HIV testing behaviors varied by group over time, independently of participants' prior lifetime STI/HIV testing behaviors. At baseline, of the 50 youths, 42 (84%) reported having been tested for STI/HIV during their lifetime (Table 1). Moreover, no baseline lifetime STI/HIV testing differences were found by group. Relative to the control group, participants in the S4E group reported an overall higher uptake of STI/HIV testing across the trial (11/25, 44% vs 13/25, 52%). Although chi-square testing (χ^2_1=0.3; *P*=.57) suggests these differences were not statistically significant, the estimated effect size (Cohen *h*=0.16, 95% CI −0.39 to 0.72) yielded a small value.

Discussion

Principal Findings

The findings suggest the feasibility of a small-scale pilot RCT and demonstrated hypothesized shifts in reducing substance use, condomless sex, and alcohol use before sex, as well as improving uptake of STI/HIV testing among a diverse sample of youth. The estimated proportion change effect sizes of our behavioral outcomes (ie, substance use, sexual risk behaviors, STI/HIV testing) and potential mechanisms of change (ie, clinician-youth communication, prevention knowledge, self-efficacy) between the S4E and enhanced usual practice control groups yielded small to large effect sizes. Drawing on previous literature, the findings provide evidence for the promise of S4E in preventing and reducing substance use and sexual risk behaviors [23,27]. Reducing substance use and sexual risk behaviors and improving uptake of STI/HIV testing have been identified as key strategies to improve the health of young people in the United States [10,15,16]. Pilot testing S4E advances the scientific knowledge on mHealth preventive interventions and has important public health implications.

We demonstrated the feasibility of measuring intermediate outcomes. The potential mechanisms underlying the hypothesized shifts in substance use, sexual risk behaviors, and STI/HIV testing among youth in the S4E group may be partially explained by improvements in clinician-youth risk communication, substance use refusal self-efficacy, and prevention knowledge. Specifically, relative to youth in the control group, S4E participants demonstrated higher levels of clinician-youth communication immediately postintervention, as well as higher levels of substance use refusal self-efficacy and prevention knowledge at 30-day follow-up. The present design precludes us from formal mediation testing. However, these findings build on previous research indicating that clinician-youth communication, self-efficacy, and STI/HIV prevention knowledge may be pathways through which S4E has an effect on substance use, sexual risk behaviors, and STI/HIV testing [39]. Future research should include at least 3 time points to allow for formal mediation analysis [52], especially because few pathways by which mHealth preventive interventions affect behavioral outcomes have been identified [53].

We demonstrated the ability to measure outcomes targeted by S4E. Evaluation of the S4E intervention suggests reductions in overall licit and illicit substance use behaviors among youth. When we separated substance use behaviors, S4E helped decrease the proportion of youths who engaged in alcohol, tobacco, or other drug use at 30-day follow-up. These findings have important public health implications because reducing substance use behaviors among the youth aligns with the nation's prevention goals and strategies to ameliorate substance use–related morbidity and mortality [15]. Furthermore, our findings are similar to those of other researcher's pilot testing preventive interventions and lend support to the promise that mHealth strategies have in preventing and reducing youth risk behaviors [24-26].

The findings that S4E demonstrated hypothesized shifts in reducing condomless sex and alcohol use before sex have important public health implications because these risk behaviors are widespread among youth [11]. Therefore, the findings that S4E participants show reductions in the proportion of youth who engage in condomless sex or alcohol use before sex have important public health implications, as these risk behaviors are linked to increased vulnerability to STIs, HIV, and unplanned pregnancy—outcomes that disproportionately affect the youth [54]. Contrary to what we hypothesized, we did not find a between-group effect on drug use before sex; however, the findings suggest a statistically significant within-group change among S4E group participants. It may be that the 30-day follow-up time period is not sufficient in duration to capture the long-term effects of S4E relative to the control group on drug use before sex. Therefore, future research should examine the effects of S4E on drug use before sex over a longer period. Importantly, findings also suggest that S4E demonstrated shifts in improving uptake of STI/HIV testing among youth. This is especially important in the realm of public health, as improving STI and HIV testing uptake among youth is a key strategy to reducing the burden of STI and HIV infection among this vulnerable population [16,54,55]. Taken together, our small-scale pilot RCT suggests the feasibility of S4E, including determining intermediate outcomes, ability to measure behavioral outcomes, as well as demonstrated hypothesized shifts in reducing substance use, sexual risk behaviors, and STI/HIV testing among youth over 30 days, which aligns with other research pilot testing technology-based interventions on these behaviors [23-28]. It is also important to note that some researchers have found that technology-based interventions have a moderate health impact on exercise over 6 weeks [56]. Thus, in general, an important future research direction is to examine both the short and long-term effects of mHealth preventive interventions. In addition, this study focused on a sample of youths seeking care in a youth-centered community health clinic. Youths who are currently not in care may be more vulnerable to substance use and sexual risk behaviors, and future research could target this population.

Our findings have important future research implications and suggest that examining the efficacy of S4E in a larger RCT may be warranted. Implementation science designs might offer other

alternatives to the traditional RCT, including stepped-wedge design and type 1 hybrid studies, which may increase the practicality of exploring intervention effects within real-life contexts [57,58]. Furthermore, given the multilevel approach of S4E, a sequential multiple assignment randomized trial (SMART) may lead to an optimally efficacious preventive strategy through the optimization of dose based on response (or lack thereof) to the intervention [59-62]. SMART is an innovative design that provides evidence for individualized decision making through adaptive interventions for the prevention of substance use and sexual risk behaviors [59,60]. In addition, our findings have important clinical implications. Our previous research establishing high feasibility and acceptability of S4E among both clinicians and youth [30,39] in conjunction with the findings of this pilot RCT endorses the implementation of strategies that support the youth-focused clinical health care workforce, especially if we are to achieve the nation's public health goals laid out by initiatives such as the Substance Abuse and Mental Health Services Administration Strategic Plan for years 2019-2023 [63], National Prevention Strategy for increasing HIV/STI testing [64], ending the HIV Epidemic in the United States by 2030 [65], and teen pregnancy objectives in healthy people by 2020 [66].

Study Limitations

Several limitations are worth noting. First, the sample in this study is not representative of the clinic population, limiting the generalizability of our findings. However, we demonstrated hypothesized shifts in prominent substance use, sexual risk behaviors, and STI/HIV testing in S4E youth, the majority of whom identified as racial and ethnic minorities. Second, the reliance on self-reported risk behavior outcomes is a limitation. Future research should consider access to medical charts as part of the study design. Third, our control group received a printed version of the S4E tobacco module content. Future research should use an attention- and time-matched control group design [67,68]. That is, an mHealth app focused on youth risk behaviors with similar intended dosage as S4E may be used in a future RCT to examine the differential effects of mHealth apps. Another limitation is that of the 211 potential participants eligible to participate in the trail; we could only reach 91 youth (43.1%). Of these, we successfully enrolled 50 youth (54.9%); however, these enrollment rates are similar to other research focused on mHealth preventive interventions with vulnerable populations [26]. The sample size of 50 is a study limitation; however, the sample size is akin to other research focused on pilot testing mHealth preventive interventions [24,26]. Finally, clinicians were not randomly assigned and provided usual practice to participants in both the experimental and control groups. Thus, there is potential for contamination between groups; however, contamination is highly unlikely given that the 2 groups are vastly different. Importantly, any potential bias would bias the findings toward the null, because clinicians could deliver S4E prevention strategies to youth in the control group.

Conclusions

In summary, this study's findings suggest the feasibility of a small-scale pilot RCT. S4E may have the potential as an mHealth strategy to reduce substance use and sexual risk–related outcomes such as STIs, HIV, and unplanned pregnancy among youth. In addition, the findings suggest that S4E demonstrated hypothesized shifts in improving uptake of STI/HIV testing, which is important for reducing the transmission and acquisition of STIs and HIV infection among youth. These findings advance scientific knowledge on mHealth preventive interventions and contribute toward improving public health through the identification of potential technology-based, youth substance use, and sexual risk behavior solutions.

Acknowledgments

All study procedures were approved by the University of Michigan Institutional Review Board (HUM00097290). A certificate of confidentiality (CC-CA-17-048-AI) was obtained from the National Institutes of Health/National Cancer Institute. This study is registered with ClinicalTrails.gov (NCT03855410). The Youth Leadership Council consisted of Ian Stewart, Erika Riano-Mojica, Bishop Warford, Franco Machado, Kiristen Hubbard, Brie Widian, Sakinah Rahman, Zakiyyah Rahman, Zeaira Chestang, and Katheryne Messer. This research was supported by pilot funding from the University of Michigan Rogel Cancer Center to David Cordova. Preparation of this manuscript was supported, in part, by grants from the National Institute of Mental Health to Torsten B Neilands (R25 MH067127), and the National Institute on Drug Abuse (R03DA04189101A1) to David Cordova.

Conflicts of Interest

None declared.

Multimedia Appendix 1
S4E youth risk behavior assessment.
[PNG File , 68 KB - mhealth_v8i3e16251_app1.png]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 3645 KB - mhealth_v8i3e16251_app2.pdf]

References

1. Centers for Disease Control and Prevention (CDC). 2019 Jul 30. STDs in Adolescents and Young Adults URL: https://www.cdc.gov/std/stats18/adolescents.htm [accessed 2019-07-12]
2. Ocfemia MC, Dunville R, Zhang T, Barrios LC, Oster AM. HIV diagnoses among persons aged 13-29 years - United States, 2010-2014. MMWR Morb Mortal Wkly Rep 2018 Feb 23;67(7):212-215 [FREE Full text] [doi: 10.15585/mmwr.mm6707a2] [Medline: 29470457]
3. Martin JA, Hamilton BE, Osterman MJ, Driscoll AK, Drake P. Births: final data for 2017. Natl Vital Stat Rep 2018 Nov;67(8):1-50. [Medline: 30707672]
4. Sedgh G, Finer LB, Bankole A, Eilers MA, Singh S. Adolescent pregnancy, birth, and abortion rates across countries: levels and recent trends. J Adolesc Health 2015 Feb;56(2):223-230 [FREE Full text] [doi: 10.1016/j.jadohealth.2014.09.007] [Medline: 25620306]
5. Centers for Disease Control and Prevention (CDC). 2018 Jul 24. STDs in Women and Infants URL: https://www.cdc.gov/std/stats17/womenandinf.htm [accessed 2019-08-15]
6. Centers for Disease Control and Prevention (CDC). 2019 Aug 6. AIDS and Opportunistic Infections URL: https://www.cdc.gov/hiv/basics/livingwithhiv/opportunisticinfections.html [accessed 2019-08-15]
7. National Conference of State Legislatures (NCSL). 2018 Oct 11. Teen Pregnancy Prevention URL: http://www.ncsl.org/research/health/teen-pregnancy-prevention.aspx [accessed 2019-08-15]
8. Ritchwood TD, Ford H, DeCoster J, Sutton M, Lochman JE. Risky sexual behavior and substance use among adolescents: a meta-analysis. Child Youth Serv Rev 2015 May;52:74-88 [FREE Full text] [doi: 10.1016/j.childyouth.2015.03.005] [Medline: 25825550]
9. Centers for Disease Control and Prevention (CDC). 2019 Aug 13. Sexual Risk Behaviors Can Lead to HIV, STDs, & Teen Pregnancy URL: https://www.cdc.gov/healthyyouth/sexualbehaviors [accessed 2019-09-10]
10. Centers for Disease Control and Prevention (CDC). 2019 Jul. Substance Use and Sexual Risk Behaviors Among Youth URL: https://www.cdc.gov/healthyyouth/substance-use/pdf/dash-substance-use-fact-sheet.pdf [accessed 2019-09-10]
11. Kann L, McManus T, Harris WA, Shanklin SL, Flint KH, Queen B, et al. Youth Risk Behavior Surveillance - United States, 2017. MMWR Surveill Summ 2018 Jun 15;67(8):1-114 [FREE Full text] [doi: 10.15585/mmwr.ss6708a1] [Medline: 29902162]
12. Cordova D, Huang S, Arzon M, Freitas D, Malcolm S, Prado G. The role of attitudes, family, peer and school on alcohol use, rule breaking and aggressive behavior in Hispanic delinquent adolescents. Open Fam Stud J 2011;4(Suppl 1-M4):38-45 [FREE Full text] [doi: 10.2174/1874922401104010038] [Medline: 22473467]
13. Clayton HB, Lowry R, August E, Jones SE. Nonmedical use of prescription drugs and sexual risk behaviors. Pediatrics 2016 Jan;137(1) [FREE Full text] [doi: 10.1542/peds.2015-2480] [Medline: 26668299]
14. Córdova D, Heinze JE, Hsieh H, Mistry R, Salas-Wright CP, Cook SH, et al. Are trajectories of a syndemic index in adolescence linked to HIV vulnerability in emerging and young adulthood? AIDS 2018 Feb 20;32(4):495-503 [FREE Full text] [doi: 10.1097/QAD.0000000000001717] [Medline: 29239889]
15. Office of Disease Prevention and Health Promotion. Healthy People 2020. 2020. Sexually Transmitted Diseases URL: https://www.healthypeople.gov/node/3523/objectives#5267 [accessed 2019-09-02]
16. Centers for Disease Control and Prevention (CDC). 2018 Apr. HIV Infection, Risk, Prevention, and Testing Behaviors Among Heterosexuals at Increased Risk for HIV Infection - National HIV Behavioral Surveillance 17 US Cities, 2016 URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-number-19.pdf [accessed 2019-09-02]
17. Centers for Disease Control and Prevention (CDC). 2013 Apr. Sexually Transmitted Infections Among Young Americans URL: https://www.cdc.gov/std/products/youth-sti-infographic.pdf [accessed 2019-09-02]
18. Buzi RS, Madanay FL, Smith PB. Integrating routine HIV testing into family planning clinics that treat adolescents and young adults. Public Health Rep 2016;131(Suppl 1):130-138 [FREE Full text] [doi: 10.1177/00333549161310S115] [Medline: 26862238]
19. Bensussen-Walls W, Saewyc EM. Teen-focused care versus adult-focused care for the high-risk pregnant adolescent: an outcomes evaluation. Public Health Nurs 2001;18(6):424-435. [doi: 10.1046/j.1525-1446.2001.00424.x] [Medline: 11737811]
20. Shafii T, Benson SK, Morrison DM. Brief motivational interviewing delivered by clinician or computer to reduce sexual risk behaviors in adolescents: acceptability study. J Med Internet Res 2019 Jul 10;21(7):e13220 [FREE Full text] [doi: 10.2196/13220] [Medline: 31293242]
21. Jemmott JB, Jemmott LS, Braverman PK, Fong GT. HIV/STD risk reduction interventions for African American and Latino adolescent girls at an adolescent medicine clinic: a randomized controlled trial. Arch Pediatr Adolesc Med 2005 May;159(5):440-449. [doi: 10.1001/archpedi.159.5.440] [Medline: 15867118]
22. World Health Organization. mHealth: New Horizons For Health Through Mobile Technologies: Second Global Survey On eHealth (Global Observatory For eHealth). Geneva: World Health Organization; 2011.
23. Bauermeister JA, Pingel ES, Jadwin-Cakmak L, Harper GW, Horvath K, Weiss G, et al. Acceptability and preliminary efficacy of a tailored online HIV/STI testing intervention for young men who have sex with men: the Get Connected! program. AIDS Behav 2015 Oct;19(10):1860-1874 [FREE Full text] [doi: 10.1007/s10461-015-1009-y] [Medline: 25638038]

24. Shrier LA, Burke PJ, Kells M, Scherer EA, Sarda V, Jonestrask C, et al. Pilot randomized trial of MOMENT, a motivational counseling-plus-ecological momentary intervention to reduce marijuana use in youth. Mhealth 2018;4:29 [FREE Full text] [doi: 10.21037/mhealth.2018.07.04] [Medline: 30148142]
25. Biello KB, Marrow E, Mimiaga MJ, Sullivan P, Hightow-Weidman L, Mayer KH. A mobile-based app (MyChoices) to increase uptake of HIV testing and pre-exposure prophylaxis by young men who have sex with men: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2019 Jan 7;8(1):e10694 [FREE Full text] [doi: 10.2196/10694] [Medline: 30617042]
26. Lelutiu-Weinberger C, Manu M, Ionescu F, Dogaru B, Kovacs T, Dorobantescu C, et al. An mHealth intervention to improve young gay and bisexual men's sexual, behavioral, and mental health in a structurally stigmatizing national context. JMIR Mhealth Uhealth 2018 Nov 14;6(11):e183 [FREE Full text] [doi: 10.2196/mhealth.9283] [Medline: 30429117]
27. Ybarra ML, Prescott T, Mustanski B, Parsons J, Bull SS. Feasibility, acceptability, and process indicators for Guy2Guy, an mHealth HIV prevention program for sexual minority adolescent boys. J Adolesc Health 2019 Sep;65(3):417-422. [doi: 10.1016/j.jadohealth.2019.04.025] [Medline: 31277991]
28. Bauermeister JA, Tingler RC, Demers M, Harper GW. Development of a tailored HIV prevention intervention for single young men who have sex with men who meet partners online: Protocol for the myDEx project. JMIR Res Protoc 2017 Jul 19;6(7):e141 [FREE Full text] [doi: 10.2196/resprot.7965] [Medline: 28724513]
29. Liu A, Coleman K, Bojan K, Serrano PA, Oyedele T, Garcia A, et al. Developing a mobile app (LYNX) to support linkage to HIV/sexually transmitted infection testing and pre-exposure prophylaxis for young men who have sex with men: protocol for a randomized controlled trial. JMIR Res Protoc 2019 Jan 25;8(1):e10659 [FREE Full text] [doi: 10.2196/10659] [Medline: 30681964]
30. Cordova D, Alers-Rojas F, Lua FM, Bauermeister J, Nurenberg R, Ovadje L, et al. The usability and acceptability of an adolescent mHealth HIV/STI and drug abuse preventive intervention in primary care. Behav Med 2018;44(1):36-47 [FREE Full text] [doi: 10.1080/08964289.2016.1189396] [Medline: 27223646]
31. Casey BJ, Getz S, Galvan A. The adolescent brain. Dev Rev 2008;28(1):62-77 [FREE Full text] [doi: 10.1016/j.dr.2007.08.003] [Medline: 18688292]
32. Centers for Disease Control and Prevention (CDC). 2019 Apr. HIV and Youth URL: https://www.cdc.gov/hiv/pdf/group/age/youth/cdc-hiv-youth.pdf [accessed 2019-08-19]
33. Forhan SE, Gottlieb SL, Sternberg MR, Xu F, Datta SD, McQuillan GM, et al. Prevalence of sexually transmitted infections among female adolescents aged 14 to 19 in the United States. Pediatrics 2009 Dec;124(6):1505-1512. [doi: 10.1542/peds.2009-0674] [Medline: 19933728]
34. Centers for Disease Control and Prevention (CDC). 2019. HIV Surveillance – Adolescents and Young Adults URL: https://www.cdc.gov/hiv/pdf/library/slidesets/cdc-hiv-surveillance-adolescents-young-adults-2017.pdf [accessed 2019-08-20]
35. Israel BA, Eng E, Schulz AJ, Parker EA. Introduction to methods for CBPR for health. In: Methods For Community-Based Participatory Research For Health. Hoboken, NJ: Wiley; 2005.
36. Cordova D, Bauermeister JA, Fessler K, Delva J, Nelson A, Nurenberg R, Youth Leadership Council. A community-engaged approach to developing an mHealth HIV/STI and drug abuse preventive intervention for primary care: a qualitative study. JMIR Mhealth Uhealth 2015 Dec 18;3(4):e106 [FREE Full text] [doi: 10.2196/mhealth.4620] [Medline: 26685288]
37. Szapocznik J, Coatsworth JD. An ecodevelopmental framework for organizing the influences on drug abuse: a developmental model of risk and protection. In: Glantz MD, Hartel CR, editors. Drug abuse: Origins & interventions. Washington, DC: American Psychological Association; 1999:331-366.
38. Zimmerman MA, Israel BA, Schulz A, Checkoway B. Further explorations in empowerment theory: an empirical analysis of psychological empowerment. Am J Commun Psychol 1992;20(6):707-727. [doi: 10.1007/bf00942234]
39. Cordova D, Lua FM, Muñoz-Velázquez J, Street K, Bauermeister JA, Fessler K, Youth Leadership Council, et al. A multilevel mHealth drug abuse and STI/HIV preventive intervention for clinic settings in the United States: a feasibility and acceptability study. PLoS One 2019;14(8):e0221508 [FREE Full text] [doi: 10.1371/journal.pone.0221508] [Medline: 31437240]
40. Schaeffer K. Pew Research Center. 2019 Aug 23. Most US Teens Who Use Cellphones Do It to Pass Time, Connect With Others, Learn New Things URL: https://www.pewresearch.org/fact-tank/2019/08/23/most-u-s-teens-who-use-cellphones-do-it-to-pass-time-connect-with-others-learn-new-things/ [accessed 2019-11-12]
41. Robertson EB, David SL, Rao SA. Preventing Drug Use Among Children And Adolescents. Bethesda, MD: National Institute on Drug Abuse; 2004.
42. Kidwell KM, Postow MA, Panageas KS. Sequential, multiple assignment, randomized trial designs in immuno-oncology research. Clin Cancer Res 2018 Feb 15;24(4):730-736 [FREE Full text] [doi: 10.1158/1078-0432.CCR-17-1355] [Medline: 28835379]
43. Oesterle TS, Hitschfeld MJ, Lineberry TW, Schneekloth TD. CRAFFT as a substance use screening instrument for adolescent psychiatry admissions. J Psychiatr Pract 2015 Jul;21(4):259-266. [doi: 10.1097/PRA.0000000000000083] [Medline: 26164051]

44. Campbell C, Lockyer J, Laidlaw T, Macleod H. Assessment of a matched-pair instrument to examine doctor-patient communication skills in practising doctors. Med Educ 2007 Feb;41(2):123-129. [doi: 10.1111/j.1365-2929.2006.02657.x] [Medline: 17269944]
45. Johnston LD, Miech RA, O'Malley PM, Bachman JG, Schulenberg JE, Patrick ME. Deep Blue. Ann Arbor, MI: Institute for Social Research, University of Michigan; 2018. Monitoring the Future National Survey Results On Drug Use, 1975-2017: Overview, Key Findings on Adolescent Drug Use URL: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/148123/Overview%202018%20FINAL%20print%201-30.pdf?sequence=1&isAllowed=y [accessed 2020-03-03]
46. Johnston LD, O'Malley RA, Miech RA, Bachman JG, Schulenberg JE. Education Resources Information Center. Ann Arbor, MI: Institute for Social Research, University of Michigan; 2018. Demographic Subgroup Trends among Adolescents in the Use of Various Licit and Illicit Drugs, 1975-2017 URL: https://eric.ed.gov/?id=ED589759 [accessed 2020-03-03]
47. Jemmott JB, Jemmott LS, Fong GT. Abstinence and safer sex HIV risk-reduction interventions for African American adolescents: a randomized controlled trial. J Am Med Assoc 1998 May 20;279(19):1529-1536. [doi: 10.1001/jama.279.19.1529] [Medline: 9605896]
48. Kendall JM. Designing a research project: randomised controlled trials and their principles. Emerg Med J 2003 Mar;20(2):164-168 [FREE Full text] [doi: 10.1136/emj.20.2.164] [Medline: 12642531]
49. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
50. IBM Corp. Armonk, NY: IBM Corp; 2016. IBM SPSS Statistics 24.0 URL: https://www.ibm.com/support/pages/release-notes-ibm-spss-statistics-240 [accessed 2020-03-03]
51. Champely S, Ekstrom C, Dalgaard P, Gill J, Weibelzahl S, Anandkumar A, et al. CRAN - R Project. 2018 Mar 3. Package 'pwr' URL: https://cran.r-project.org/web/packages/pwr/pwr.pdf [accessed 2020-03-03]
52. Carper MM, Makover HB, Kendall PC. Future directions for the examination of mediators of treatment outcomes in youth. J Clin Child Adolesc Psychol 2018;47(2):345-356 [FREE Full text] [doi: 10.1080/15374416.2017.1359786] [Medline: 28841335]
53. de Korte E, Wiezer N, Roozeboom MB, Vink P, Kraaij W. Behavior change techniques in mHealth apps for the mental and physical health of employees: systematic assessment. JMIR Mhealth Uhealth 2018 Oct 3;6(10):e167 [FREE Full text] [doi: 10.2196/mhealth.6363] [Medline: 30282621]
54. Centers for Disease Control and Prevention (CDC). 2013. Sexually Transmitted Infections among Young Americans Infographic URL: https://www.cdc.gov/nchhstp/newsroom/2013/sam-2013.html [accessed 2019-08-21]
55. Conserve DF, Jennings L, Aguiar C, Shin G, Handler L, Maman S. Systematic review of mobile health behavioural interventions to improve uptake of HIV testing for vulnerable and key populations. J Telemed Telecare 2017 Feb;23(2):347-359 [FREE Full text] [doi: 10.1177/1357633X16639186] [Medline: 27056905]
56. Howe KB, Suharlim C, Ueda P, Howe D, Kawachi I, Rimm EB. Gotta catch'em all! Pokémon GO and physical activity among young adults: difference in differences study. Br Med J 2016 Dec 13;355:i6270 [FREE Full text] [doi: 10.1136/bmj.i6270] [Medline: 27965211]
57. Bernet AC, Willens DE, Bauer MS. Effectiveness-implementation hybrid designs: implications for quality improvement science. Implementation Sci 2013;8(Suppl 1):S2. [doi: 10.1186/1748-5908-8-S1-S2]
58. Hemming K, Taljaard M, McKenzie JE, Hooper R, Copas A, Thompson JA, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. Br Med J 2018 Nov 9;363:k1614 [FREE Full text] [doi: 10.1136/bmj.k1614] [Medline: 30413417]
59. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A 'SMART' design for building individualized treatment sequences. Annu Rev Clin Psychol 2012;8:21-48 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032511-143152] [Medline: 22224838]
60. Kidwell KM, Hyde LW. Adaptive Interventions and SMART Designs: application to child behavior research in a community setting. Am J Eval 2016 Sep;37(3):344-363 [FREE Full text] [doi: 10.1177/1098214015617013] [Medline: 28239254]
61. Almirall D, Nahum-Shani I, Sherwood NE, Murphy SA. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. Transl Behav Med 2014 Sep;4(3):260-274 [FREE Full text] [doi: 10.1007/s13142-014-0265-0] [Medline: 25264466]
62. Vock DM, Almirall D. Wiley Online Library. 2018 Aug. Wiley StatsRef: Statistics Reference Online - Sequential Multiple Assignment Randomized Trial (SMART) URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat08073 [accessed 2020-03-03]
63. Substance Abuse and Mental Health Services Administration (SAMHSA). 2018 Nov 2. SAMHSA Strategic Plan FY2019-FY2023 URL: https://www.samhsa.gov/about-us/strategic-plan [accessed 2019-09-12]
64. US Department of Health and Human Services. HIV.gov. 2015 Aug 2. National HIV/AIDS Strategy: Updated to 2020. 5 Major Changes Since 2010 URL: https://files.hiv.gov/s3fs-public/nhas-update-5-things.pdf [accessed 2019-09-20]
65. UNAIDS. The Joint United Nations Programme on HIV/AIDS (UNAIDS). 2014 Oct. 90-90-90: An Ambitious Treatment Target to Help End the AIDS Epidemic URL: https://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf [accessed 2019-09-15]
66. US Department of Health and Human Services. Healthy People 2020. 2020. Family Planning URL: https://www.healthypeople.gov/2020/topics-objectives/topic/family-planning/objectives [accessed 2019-09-02]

67. Rose S, Laan MJ. Why match? Investigating matched case-control study designs with causal effect estimation. Int J Biostat 2009 Jan 6;5(1):Article 1 [FREE Full text] [doi: 10.2202/1557-4679.1127] [Medline: 20231866]
68. Aycock DM, Hayat MJ, Helvig A, Dunbar SB, Clark PC. Essential considerations in developing attention control groups in behavioral research. Res Nurs Health 2018 Jun;41(3):320-328. [doi: 10.1002/nur.21870] [Medline: 29906317]

Abbreviations

ATOD: alcohol, tobacco, other drug use
CBPR: community-based participatory research
mHealth: mobile Health
MPI: Matched Pair Instrument
RCT: randomized controlled trial
S4E: Storytelling For Empowerment
SMART: Sequential Multiple Assignment Randomized Trial
STI: sexually transmitted infection

Edited by G Eysenbach; submitted 19.09.19; peer-reviewed by L Shrier, C Yang, Z Ma; comments to author 09.10.19; revised version received 04.12.19; accepted 10.02.20; published 17.03.20.

Please cite as:
Cordova D, Munoz-Velazquez J, Mendoza Lua F, Fessler K, Warner S, Delva J, Adelman N, Youth Leadership Council, Fernandez A, Bauermeister J
Pilot Study of a Multilevel Mobile Health App for Substance Use, Sexual Risk Behaviors, and Testing for Sexually Transmitted Infections and HIV Among Youth: Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(3):e16251
URL: https://mhealth.jmir.org/2020/3/e16251
doi:10.2196/16251
PMID:32181747

Original Paper

A Mobile Health Intervention for Mental Health Promotion Among University Students: Randomized Controlled Trial

Marcus Bendtsen[1], PhD; Ulrika Müssener[1], PhD; Catharina Linderoth[1], BSc; Kristin Thomas[1], PhD

Linköping University, Linköping, Sweden

Corresponding Author:
Kristin Thomas, PhD
Linköping University
581 83 Linköping
Linköping
Sweden
Phone: 46 13282546
Email: kristin.thomas@liu.se

Abstract

Background: High positive mental health, including the ability to cope with the normal stresses of life, work productively, and be able to contribute to one's community, has been associated with various health outcomes. The role of positive mental health is therefore increasingly recognized in national mental health promotion programs and policies. Mobile health (mHealth) interventions could be a cost-effective way to disseminate positive psychological interventions to the general population.

Objective: The aim of this study was to estimate the effect of a fully automated mHealth intervention on positive mental health, and anxiety and depression symptomology among Swedish university students using a randomized controlled trial design.

Methods: A 2-arm, single-blind (researchers), parallel-groups randomized controlled trial with an mHealth positive psychology program intervention group and a relevant online mental health information control group was employed to estimate the effect of the novel intervention. Participants were recruited using digital advertising through student health care centers in Sweden. Inclusion criteria were (1) university students, (2) able to read and understand Swedish, (3) and have access to a mobile phone. Exclusion criteria were high positive mental health, as assessed by the Mental Health Continuum Short Form (MHC-SF), or high depression and anxiety symptomology, as assessed by the Hospital Anxiety Depression Scale (HADS). The primary outcome was positive mental health (MHC-SF), and the secondary outcomes were depression and anxiety symptomatology (HADS). The subscales of MHC-SF were also analyzed as exploratory outcomes. Outcomes were measured 3 months after randomization through questionnaires completed on the participants' mobile phones.

Results: A total of 654 participants (median age 25 years), including 510 (78.0%) identifying as female, were randomized to either the intervention (n=348) or control group (n=306). At follow-up, positive mental health was significantly higher in the intervention group compared with the control group (incidence rate ratio [IRR]=1.067, 95% CI 1.024-1.112, *P*=.002). For both depression and anxiety symptomatology, the intervention group showed significantly lower scores at follow-up compared with the control group (depression: IRR=0.820, 95% CI 0.714-0.942, *P*=.005; anxiety: IRR=0.899, 95% CI 0.840-0.962, *P*=.002). Follow-up rates were lower than expected (58.3% for primary outcomes and 52.3% for secondary outcomes); however, attrition analyses did not identify any systematic attrition with respect to baseline variables.

Conclusions: The mHealth intervention was estimated to be superior to usual care in increasing positive mental health among university students. A protective effect of the intervention was also found on depressive and anxiety symptoms. These findings demonstrate the feasibility of using an automated mobile phone format to enhance positive mental health, which offers promise for the use of mHealth solutions in public mental health promotion.

Trial Registration: International Standard Randomized Controlled Trial Registry ISRCTN54748632; http://www.isrctn.com/ISRCTN54748632

(JMIR Mhealth Uhealth 2020;8(3):e17208) doi:10.2196/17208

KEYWORDS

mHealth; positive mental health; university students; randomized controlled trial

Introduction

Background

A substantial body of research has shown a link between high positive mental health and decreased risk of disease [1-5], decreased risk of mental illness [6-9], and increased longevity [3,10]. The promotion of positive mental health among the general population has recently been stressed as the most important goal for the public mental health agenda in Europe [11]. This stems from longitudinal research suggesting a protective effect of positive mental health on mental health problems [9,12].

"Positive mental health" has been defined to encompass feelings of happiness and satisfaction with life (emotional well-being), positive individual functioning regarding self-realization (psychological well-being), and positive societal functioning (social well-being). Furthermore, the two-continua model holds that mental illness and mental health are related but distinct dimensions [13]. The working theory of positive psychology interventions (PPIs) is that elevated positive emotions, thoughts, and behaviors will lead to increased positive mental health [14,15]. PPIs strive to increase the frequency of positive emotions, thoughts, and behaviors through exercises. For instance, to increase positive thinking of gratitude, individuals are asked "to think about three things that you are grateful for today" [5,15]. Although there is evidence to support the efficiency of PPIs on positive mental health for both healthy and clinical populations, more research is needed.

Two meta-analyses of PPIs reported relatively small effect sizes [5,15], which may be due to the composition of the included interventions. Most studies included in the meta-analyses examined the effect of individual exercises targeting only one aspect of positive mental health (eg, using a gratitude journal to increase positive thinking). However, as positive mental health is a multilayered construct [13], a multicomponent intervention taking into account emotional, social, and psychological well-being may be more effective. A systematic review and meta-analysis on multicomponent PPIs found supporting evidence in terms of positive mental health and depression, and potentially anxiety and stress. However, the authors concluded that larger and more rigorous studies are needed to move the research field forward, for instance through sufficiently powered trials with transparent methodological reporting [16]. In addition, a limited number of studies included in the aforementioned reviews were delivered through mobile health (mHealth) interventions, an otherwise fast-growing practice and research field.

mHealth interventions could potentially be a cost-effective means to disseminate PPIs to a large population [17,18]. There are now countless mobile apps commercially available that target positive mental health among the general population; however, the majority of these apps lack experimental evidence, are not theory-based, and have not been scientifically evaluated [19]. A review on the effect of digital interventions (eg, mobile apps) on mental health showed a small to medium effect, suggesting that mental health problems decreased while positive mental health increased among those with access to the interventions. However, the overall quality of the body of evidence in the review was low due to several concerns regarding risk of bias [20]. Another review summarized the evidence for theory-driven and evidence-based mental health eResources (eg, website or mobile apps) and only found one randomized controlled trial, suggesting a lack of valid evidence. The authors concluded that eResources for mental health have the potential to be widely effective, but that more rigorous studies are needed to clarify the benefits [21].

Objectives

The aim of this study was to estimate the effect of a fully automated mHealth intervention on positive mental health and anxiety and depression symptomology among Swedish university students using a randomized controlled trial design.

The primary hypothesis was that positive mental health will differ among groups at 3 months postrandomization, with those having access to the novel mHealth intervention reporting higher scores on the Mental Health Continuum Short Form (MHC-SF). The secondary hypotheses were: (1) depression and anxiety symptomology will differ among groups at 3 months postrandomization, with those having access to the novel mHealth intervention reporting lower scores on the subscales of the Hospital Anxiety Depression Scale (HADS); and (2) emotional, social, and psychological well-being will differ among groups at 3 months postrandomization, with those having access to the novel mHealth intervention reporting higher scores on the subscales of the MHC-SF.

Methods

Trial Design

This trial was prospectively registered with the International Standard Randomized Controlled Trial Registry (ISRCTN54748632) and a trial protocol was made available prior to trial commencement [22]. The study received ethical approval from the Regional Ethical Review Board of Linköping University, Sweden (Dnr. 2018/519-32).

A 2-arm, single-blind (researchers), parallel-groups randomized controlled trial (1:1) was employed to estimate the effect of the novel mHealth intervention. Participants were allocated to either an intervention group (mHealth program) or control group (treatment as usual). Prior to trial commencement, but after trial registration and publication of the protocol, it was decided to remove age restrictions in the eligibility criteria. This was done so that participants would more accurately represent all Swedish university students and not only young adults.

Participants

Inclusion criteria were (1) university students, (2) able to read and understand Swedish, (3) and have access to a mobile phone. The exclusion criterion was high positive mental health defined as a score of 70 or more on the MHC-SF [23]. As the intervention was not designed to treat mental health problems, a second exclusion criterion was depression and anxiety symptomatology defined as a score of greater than or equal to 10 on both subscales of the HADS [24]. Individuals excluded due to a high HADS score were encouraged to seek help and

were provided information on where to receive support (contact information of their local student health center, primary care center, or governmental national health website).

Recruitment of students was carried out at 15 universities in Sweden, which lasted between October 8, 2018 and April 30, 2019. Recruitment was achieved through digital advertising, including email, university websites, student health care center websites, and learning management systems used by the universities. The advertisement included information on the study aims, confidentiality, and trial design. Students indicated their interest in taking part in the trial by texting a dedicated telephone number included in the advertisement material. The students then received a text message response with a link to the informed consent form and completed an online baseline questionnaire on their mobile phones. Eligibility was determined from responses to the baseline questionnaire, and eligible participants were automatically randomized to either the intervention or control group. Participants were given information on which group they had been allocated to.

Intervention

The intervention was a fully automated mHealth positive psychology multicomponent program. The program was based on theories and empirical evidence from the positive psychology research field [13,25] and aimed to enhance users' positive mental health. The program encompassed information about well-being, validated self-help exercises, brief tips, self-monitoring, and personalized feedback. Text messages were sent to users throughout the program, with an average of one text message per day, and included text and links to interactive exercises and further reading. The program ran for 10 weeks, with a new theme introduced each week. The themes used have been shown to be important for positive mental health, and included gratitude, savoring, positive emotions, personal strengths, positive relations, social environment, health behaviors, optimism, and goal setting. In the final week, users were recommended to reflect on the program, for instance by writing down any lessons learnt. Details of the intervention can be found elsewhere (Multimedia Appendix 1) [22].

Participants allocated to the control group were informed of their allocation status via a text message. The text message included contact details to their local student health center, primary care center, and governmental national health website. At the time of the trial, this was considered treatment as usual.

Outcomes

As the primary outcome, positive mental health was measured using the 14-item MHC-SF [23], in which higher scores indicate greater emotional, social, and psychological well-being (range 0-84). Secondary outcomes included depression and anxiety symptomatology, measured as the score on corresponding subscales of the HADS [24]. Each subscale consists of 7 items for a total of 14 items. Item scores are calculated into a total scale score for anxiety (range 0-21) and depression (range 0-21), with higher scores indicating higher depression and anxiety. We further measured emotional, social, and psychological well-being as exploratory outcomes based on the mean score of each subscale of the MHC-SF.

A SPIRIT [26] checklist depicting study procedures and measurements is presented in Table 1. At baseline, the participants' age, gender, and social status were recorded, along with primary, secondary, and exploratory outcomes. Two face-valid mediator items, aimed at measuring the frequency of positive thoughts and emotions, were also determined at baseline according to responses to the following questions: ''During the last week, to what extent have you experienced positive thoughts?" and "During the last week, to what extent have you experienced positive emotions?" Participants were asked to rate their response on a scale ranging from 1 (''not at all'') to 10 (''to a very high extent'').

Five weeks after randomization, participants were sent a text message with a link to the two mediator items. Three months after randomization, participants were sent a text message with a link to the follow-up questionnaire. The follow-up questionnaire explored primary, secondary, and exploratory outcomes (MHC-SF and HADS).

Participants that did not respond to the initial follow-up attempt were sent up to 4 reminders 2 days apart. Those who had still not responded were called by telephone (maximum 3 call attempts).

Table 1. SPIRIT checklist depicting study procedures, measurements, and timeline.

	Study Period				
	Enrollment	Allocation	Post-allocation		Follow-up
		0	Week 5	Week 10	3 months
Enrollment					
Eligibility screen		X			
Informed consent	X				
Interventions					
Text message intervention		X	X	X	
Referred to sources of mental health information		X	X	X	
Assessments					
Baseline demographics		X			
MHC-SF[a]		X			X
Mediators		X	X		X
HADS[b]		X			X

[a]MHC-SF: Mental Health Continuum Short Form.

[b]HADS: Hospital Anxiety Depression Scale.

Sample Size

A power analysis was conducted to determine the necessary number of participants to recruit for the study. To detect a standardized effect size of 0.3, the average score in the intervention group should exceed scores of 62% of the control group; hence, a total of 352 participants was expected to be required. The calculations were performed assuming an 80% chance of detecting a difference at a significance level of alpha=.05 (two-tailed). Assuming that 70% of the participants would respond to the follow-up questionnaire, it was deemed necessary to recruit a total of 503 participants.

Randomization and Blinding

After completing the online baseline questionnaire, participants were randomly allocated a number 0 or 1 with equal probability using Java's built-in random number generator (java.util. Random). Participants with a 0 were allocated to the control group, and participants with a 1 were allocated to the intervention group. Generation of the randomization sequence was therefore fully computerized, and allocation was concealed from participants and research team members.

Statistical Analysis

All analyses conformed to the prespecified statistical analysis plan in the trial protocol [22]. For Hypothesis 1, the primary outcome (score of the MHC-SF) was analyzed. This score represents a discrete measure that may be skewed; thus, we regressed this outcome on group allocation and baseline variables using negative binomial regression. Hypothesis 2 was investigated by analyzing secondary outcome measures (subscales of the HADS). This score is also a discrete measure that may be skewed, and was regressed against group allocation and baseline variables using negative binomial regression. Hypothesis 3 was investigated by analyzing the subscales of the MHC-SF: emotional well-being, social well-being, and psychological well-being. These are mean scores from Likert scale items, which should tend toward normality owing to the law of large numbers. Therefore, we regressed the individual scores against group allocation and baseline variables using normal linear regression.

Analytical Approaches

Analyses were performed under the intention-to-treat principle, including all randomized individuals, and no imputations were made for missing values. Missing outcome data were initially handled by a complete-case analysis, which assumed that data were missing completely at random (MCAR). If data are systematically missing, this may indicate that early responders differ from late responders, and in extension that late responders are more similar to nonresponders. We therefore explored the plausibility of the MCAR assumption by regressing the primary outcomes on the number of follow-up attempts needed before a response was recorded. To further explore the MCAR assumption, attrition was investigated among study groups by comparing baseline characteristics between those who did and did not respond to the follow-up questionnaire.

Both adjusted and unadjusted models were designed when investigating all hypotheses; however, it was decided a priori to primarily use adjusted models [22]. For all models, the coefficients of interest were assessed for statistical significance using a null hypothesis testing approach, where all tests were two-tailed at a significance level of alpha=.05.

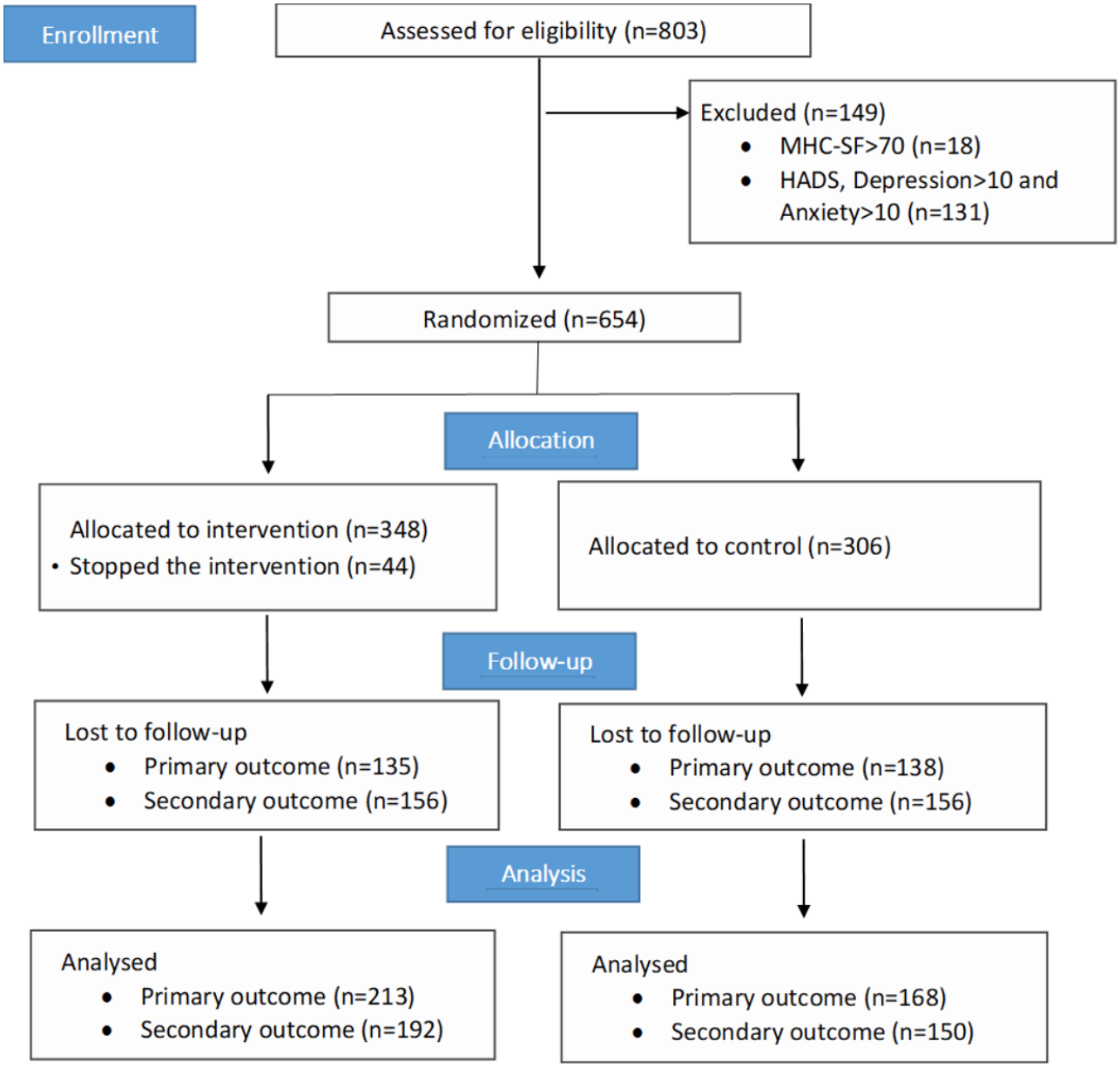
Effect modification tests were performed to assess if any of the baseline characteristics moderated the effect of the intervention.

Results

Recruitment

Figure 1 depicts a CONSORT diagram of the trial participant flow. A total of 654 participants were randomized: 348 (53.2%) to the intervention group and 306 (46.8%) to the control group. In the intervention group, 340 (79.2%) participants received all messages, and the remaining 44 individuals decided to stop the intervention before completion. All participants were contacted at follow-up regardless of adherence to the intervention. A total of 381 (58.3%) participants were included in the analysis of the primary outcomes, and 342 (52.3%) were included in the analysis of secondary outcomes. No participants explicitly requested to be removed from the trial.

Figure 1. CONSORT diagram of trial participant flow. HADS: Hospital Anxiety Depression Scale; MHC-SF: Mental Health Continuum Short Form.



Baseline Data

Table 2 summarizes the baseline characteristics of the recruited participants. There were no statistically significantly differences between the intervention and control groups with respect to any of the measured baseline characteristics. Notably, the great majority of participants identified as female.

Table 2. Baseline characteristics for both study groups.

Variable	Intervention (n=348)	Control (n=306)	*P* value
Gender, n (%)			.42[a]
Female	277 (79.6)	233 (76.1)	
Male	69 (19.8)	69 (22.5)	
Other	2 (0.6)	4 (1.3)	
Age (years), median (IQR[b])	25 (22-29)	26 (22-30)	.28[c]
Marital status, n (%)			.73[a]
No partner; no children at home	170 (48.9)	149 (48.7)	
No partner; children at home	7 (2.0)	6 (2.0)	
Partner; no children at home	95 (27.3)	72 (23.5)	
Partner; children at home	46 (13.2)	47 (15.4)	
Partner but not living together	30 (8.6)	32 (10.5)	
Positive thoughts, median (IQR)	6 (4-7)	6 (4-7)	.43[c]
Positive emotions, median (IQR)	5 (4-7)	5 (4-7)	.80[c]
Anxiety[d], median (IQR)	12 (10-15)	12 (10-14)	.58[c]
Depression[e], median (IQR)	6 (4-8)	6 (4-8)	.93[c]
Total well-being[e], median (IQR)	47.5 (40-55)	50 (40.25-57)	.14[c]
Emotional well-being[f], mean (SD)	3.79 (0.89)	3.77 (0.92)	.78[g]
Social well-being[f], mean (SD)	2.93 (0.92)	3.06 (0.89)	.07[g]
Psychological well-being[f], mean (SD)	3.59 (0.91)	3.67 (0.87)	.24[g]

[a]Fisher exact test.

[b]IQR: interquartile range.

[c]Mann-Whitney *U* test.

[d]Hospital Anxiety Depression Scale.

[e]Mental Health Continuum Short Form.

[f]Subscales of the Mental Health Continuum Short Form calculated as means of responses to each respective subset of questions.

[g]Student *t* test.

Outcomes

At 3 months after randomization, primary outcome data were collected from 213 (61.2%) participants in the intervention group and 168 (54.9%) participants in the control group. Secondary outcome data were collected from 192 (55.2%) participants in the intervention group and from 150 (49.0%) participants in the control group. These data were used to investigate the trial hypotheses according to the statistical analysis plan. The results are summarized in Table 3.

At the 3-month follow-up, positive mental health measured by the MHC-SF was significantly higher in the intervention group compared with that of the control group, which supported Hypothesis 1. Cronbach alpha was .91 including all items of MHC-SF, indicating high reliability of the measure. In addition, the scores of depression and anxiety symptoms (subscales of HADS) were both significantly lower in the intervention group compared with those of the control group, supporting Hypothesis 2. Cronbach alpha for the items included in the anxiety subscale was .81, and was .83 for the items included in the depression subscale, indicating high reliability of both measures. The subscales of the MHC-SF were all significantly higher in the intervention group compared with those of the control group. Emotional well-being, social well-being, and psychological well-being scores were all in the anticipated direction supporting Hypothesis 3. Cronbach alpha indicated high reliability for all three subscales at .83, .79, and .85 for emotional, social, and psychological well-being, respectively.

Table 3. Summary of trial hypotheses analyses at the 3-month follow-up assessment.

Outcome	Follow-up	Difference[a]	Unadjusted model		Adjusted model[b]	
			Regression coefficient[c] (95% CI)	*P* value	Regression coefficient (95% CI)	*P* value
Primary outcome						
Total well-being, median (IQR[d])						
Intervention (n=213)	56 (47 to 65)	7 (1 to 13)	1.035 (0.985-1.087)	.17	1.067 (1.024-1.112)	.002
Control (n=168)	56 (42 to 64.25)	3 (–4 to 10.25)				
Secondary outcomes						
Depression, median (IQR)						
Intervention (n=192)	4 (2 to 6)	–2 (–4 to 0)	0.817 (0.699-0.954)	.01	0.820 (0.714-0.942)	.005
Control (n=150)	4 (2 to 8)	–1 (–3 to –1)				
Anxiety, median (IQR)						
Intervention (n=192)	9 (7 to 12)	–2 (–4.25 to 0)	0.926 (0.854-1.004)	.06	0.899 (0.840-0.962)	.002
Control (n=150)	10 (8 to 13)	–1 (–3 to –1)				
Exploratory outcomes						
Emotional well-being, mean (SD)						
Intervention (n=213)	4.23 (0.88)	0.45 (0.86)	0.152 (–0.037-0.341)	.11	0.222 (0.062-0.383)	.007
Control (n=168)	4.08 (0.99)	0.20 (0.93)				
Social well-being, mean (SD)						
Intervention (n=213)	3.47 (1.03)	0.51 (0.89)	0.082 (–0.136-0.300)	.46	0.203 (0.021-0.385)	.03
Control (n=168)	3.39 (1.13)	0.25 (0.98)				
Psychological well-being, mean (SD)						
Intervention (n=213)	4.20 (0.99)	0.60 (0.90)	0.166 (–0.040-0.372)	.11	0.272 (0.093-0.451)	.003
Control (n=168)	4.04 (1.05)	0.28 (0.99)				

[a]Difference between follow-up measurement and baseline measurement.

[b]Adjusted for outcome measure, gender, age, marital status, and mediators.

[c]Incidence rate ratio for group by negative binomial regression for total well-being, depression, and anxiety; linear coefficient for emotional well-being, social well-being, and psychological well-being.

[d]IQR: interquartile range.

Sensitivity Analyses

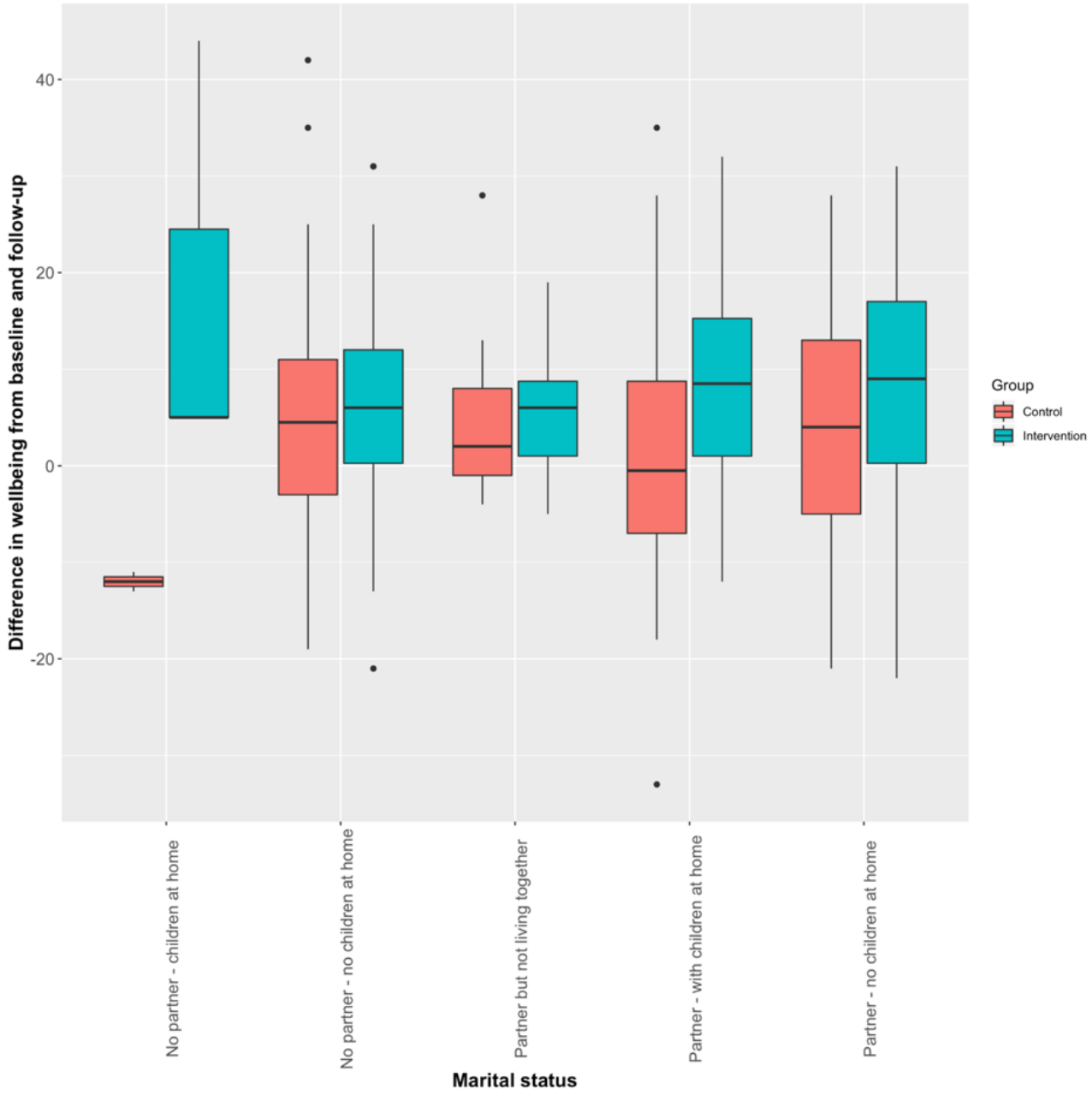
Sensitivity analyses were carried out for the primary and secondary hypotheses in which the missing outcomes were set to baseline values. No large differences in direction or statistical significance were found for the three analyses (total well-being incidence rate ratio [IRR]=1.052, 95% CI 1.024-1.08, *P*<.001; depression IRR=0.878, 95% CI 0.817-0.943, *P*<.001; anxiety IRR=0.936, 95% CI 0.894-0.981, *P*=.006).

Effect Modification Analyses

Effect modification models were explored for the primary outcome. All variables measured at baseline were considered in separate models, and were compared with the noninteraction model using likelihood ratio tests. There was weak evidence for an interaction effect with respect to age (*P*=.03) and marital status (*P*=.03).

Exploring the interaction with marital status through a box plot (Figure 2) revealed that it was the marital status category "No partner; children at home" that was driving the modification effect. Among the participants analyzed, only 5 responded with this option, and removing them from the analysis removed the interaction effect for both marital status and age. Overall, any interaction between group allocation, age, and marital status was not considered to be strongly supported by the data.

Figure 2. Interaction between group allocation and marital status.



Attrition Analyses

We used two approaches to explore the MCAR assumption: (1) response/nonresponse at follow-up was regressed against baseline characteristics, and (2) primary outcome was regressed against the number of attempts required to reach the respondent to collect follow-up data. The second approach assumes that nonresponders are actually late responders; thus, if there is an association between follow-up attempts with primary outcome, then nonresponders may be systematically different from responders. However, there was no significant association between number of follow-up attempts and response, suggesting that there was no systematic difference between early and late responders.

When regressing response/nonresponse against baseline characteristics, we found that age was potentially associated with response, as the odds ratio was 1.03 (95% CI 1.008-1.053, *P*=.008), suggesting that older participants were more likely to respond to follow-up. Locally weighted scatterplot smoothing confirmed this association (Figure 3), which was mainly driven by participants aged 40 or more, 40/54 (74.1%) of whom responded to follow-up, compared to 338/597 (56.6%) among those under 40 years of age. No such association between age and response was observed among participants under 40 years old (Figure 4).

Figure 3. Locally weighted scatterplot smoothing: age against proportion of participants responding to follow-up for all participants.

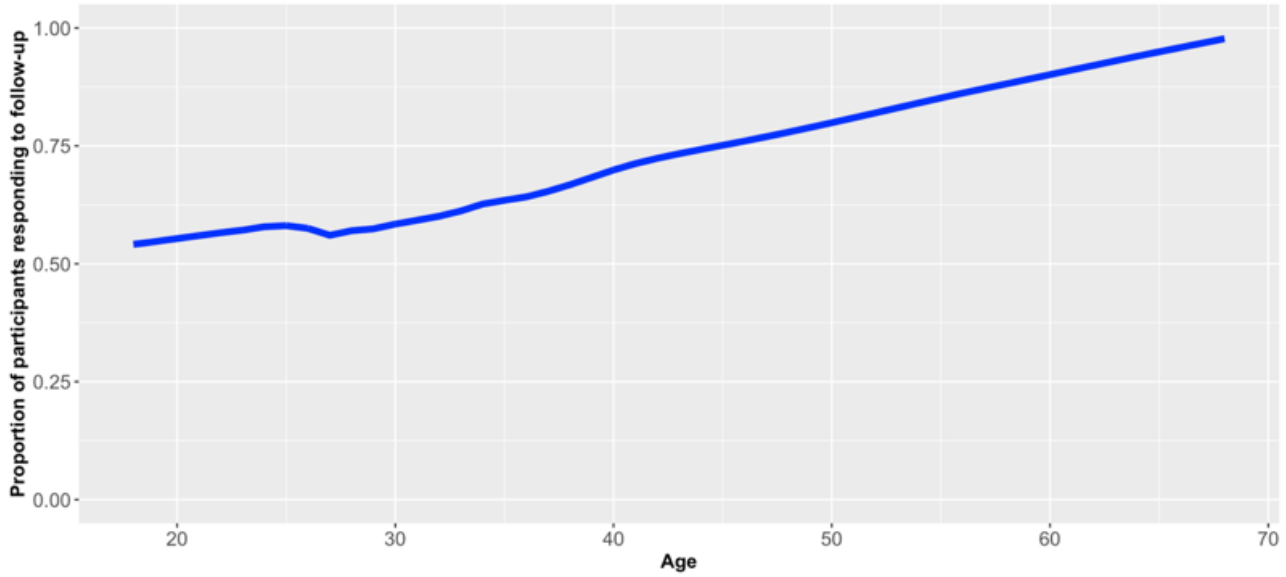
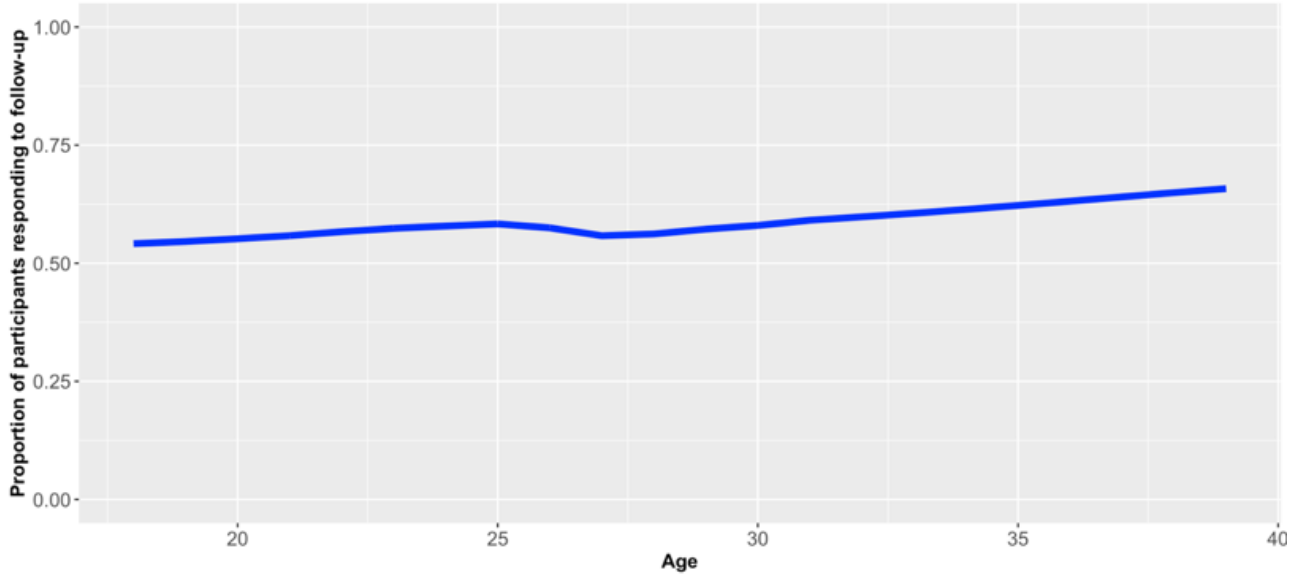


Figure 4. Locally weighted scatterplot smoothing: age against proportion of participants responding to follow-up, including only participants aged less than 40 years at baseline.



Discussion

Principal Findings

We examined whether a fully automated mHealth intervention was effective in increasing positive mental health among university students. The results suggested that the intervention may be superior to usual care in increasing positive mental health for this group (MHC-SF: IRR 1.067, 95% CI 1.024-1.112, *P*=.002). In addition, the results indicated a protective effect of the intervention on depressive and anxiety symptoms.

In general, our findings confirm previous research on the effect of PPIs on mental health. Reviews investigating PPIs delivered via face-to-face and self-help (not mHealth) modalities have shown an increase in positive mental health and enabled individuals to manage mental health problems (eg, based on decreased anxiety and worry scores) [5,15]. Interestingly, studies using digital-based interventions (eg, DVD, VCR, or Web-based) showed a comparable effect on mental health, supporting the viability of digital interventions in this area [27,28]. For example, a recent study found that an intervention encompassing self-help book and email support successfully increased positive mental health and decreased depression among adults [29]. Our findings add to this growing body of evidence, and further demonstrate the effect and feasibility of delivering PPIs in a fully automated mobile phone format.

Furthermore, our findings show that a relatively inclusive program (10 weeks including 9 positive psychology themes) was acceptable for participants to engage with, as a strong majority of participants completed the program. Research on multicomponent PPIs has shown that these integral programs influence positive mental health and depression, and potentially anxiety and stress [16]. Our study adds to this body of research showing that a quite comprehensive program (including several activities that targeted several well-being components) was

feasible and acceptable among the target group and was effectively delivered via mobile phones.

Several factors that can moderate the effect of PPIs have been proposed, including the duration of an intervention, baseline affect states of participants (eg, healthy vs subclinical populations), and recruitment methods (self-referral vs referred by a health care practitioner). An early review on PPIs delivered in group or self-administered contexts indicated that the optimal duration of PPIs was about 8 weeks [15]. Our findings provide support for the acceptability and feasibility of longer-duration PPIs. In future research, it would be interesting to investigate further dose-response and person-activity fit aspects of PPIs; that is, the effect of the variety, frequency, and tailoring of activities on mental health outcomes. PPIs tend to be more effective among subclinical populations (eg, moderate anxiety and depressiveness). Along similar lines, although our intervention was developed as a preventative measure targeting the healthy population, the findings indicate a protective effect on anxiety and depressive symptomatology for this group. Furthermore, there is controversy in the literature as to whether PPIs are more effective among self-referral groups or when they are disseminated through health care referral routes; however, self-referral seems to be more effective when considering only healthy populations [4,15]. From a public health perspective, recruitment via self-referral to self-help PPIs can offer cost-effective mental health promotion tools to reach large target groups.

Finally, the participants in our study were mostly women. Although our analyses did not indicate bias regarding gender, the generalizability of the findings could be limited. Previous research investigating PPIs that have used similar recruitment methods (self-selected) reported similar rates (about 85% women) [29]. This could indicate a greater interest and willingness among women to take part in well-being studies. However, as positive mental health does not necessarily differ between men and women, future studies could benefit from identifying recruitment strategies that can reach men to a greater extent.

Limitations

The trial design does not allow us to isolate specific themes of the program to estimate their individual effects. That is, we cannot identify if a specific theme and its accompanying exercises (eg, practicing gratitude) increased positive mental health to a greater extent than other themes. However, the literature suggests that single-component interventions have smaller effect sizes compared to multicomponent interventions [5,29], indicating that it would be more effective to invest in the latter to increase well-being among the general population. Future studies with alternative designs are needed to investigate how different components of the program, or combination of components, contribute to mental health outcomes.

A prominent limitation of this trial is the risk of attrition bias due to low follow-up rates. Despite the lack of strong evidence against MCAR, the assumption cannot be formally tested, and therefore the presented results should be interpreted under this limitation. We designed the trial under the assumption of achieving similar follow-up rates as in previous trials on the same target group (reaching as high as 90% of students [30,31]); however, we could not achieve the same retention rate in this trial. One factor causing this increased attrition was the design of the MHC-SF, which is hard to complete over the telephone given that it is relatively long and difficult to communicate verbally. Although the use of validated instruments such as MCH-SF and HADS is a strength of this trial, future research on PPIs targeting university students should consider other item sets for measuring outcomes to increase retention.

We did not include an active control in this trial, mainly with a view to estimate the total effect of the intervention compared to minimal contact, but also since there are no other available effective interventions that could be offered at this scale and through digital means. There are limitations to this approach, as it makes blinding difficult and may therefore result in performance bias. To reduce the risk of performance bias, we ensured that all participants were treated equally by automating all processes and not having any interaction between the research personnel and participants. However, one factor increasing the risk of detection bias stems from the use of telephone follow-up for nonresponders, as participants may have revealed to the interviewer which group they belonged to. We decided to go ahead with telephone follow-ups as the benefit in terms of reduced attrition bias was believed to outweigh the risk of detection bias, and interviewers were instructed to avoid discussions about group allocation.

Participants in this trial were recruited naturally through passive advertisement in printed and digital media provided by the student health care centers. This closely mimics the way that individuals would come in contact with the intervention had it been disseminated outside the trial setting. Thus, the participants recruited are comparable to participants who would use the intervention in a real-world setting. The results herein can be considered generalizable to a wider context of university students; however, the limitations discussed above should be taken into consideration before any decision for dissemination is made.

Conclusions

An mHealth intervention based on theories and empirical evidence from the positive psychology research field [16,25] was estimated to have a superior effect on positive mental health compared to usual care (MHC-SF: IRR 1.067, 95% CI 1.024-1.112, P=.002). In addition, a protective effect of the intervention was found on depressive and anxiety symptoms. These findings demonstrate the feasibility of using an automated mobile phone format to enhance positive mental health.

Disseminating mHealth PPIs could have significant public health benefits. It has been estimated that even small improvements in positive mental health in the general population could yield large preventive effects on psychopathology [32]. Previous studies have shown that targeting PPIs among people with low to moderate well-being not only increases positive mental health but can also prevent future anxiety and depression [7,8]. Our findings support the notion that mHealth solutions can be used in public mental health promotion.

Future research should focus on estimating the effect of PPI mHealth interventions among other societal groups, including studies with the elderly, minorities, and general population. Dismantling of the content of the mHealth intervention to identify which combinations of themes show the greatest effect can guide future developments of PPIs, such that even more effective interventions may be developed. Finally, creating strategies to implement and disseminate PPI mHealth interventions to a general audience will also be necessary to maximize the societal benefit of these interventions.

Acknowledgments

The authors would like to thank Preben Bendtsen, Nadine Karlsson, and Matti Leijon for their valuable input when developing the intervention. The project was funded by Linköping University.

Authors' Contributions

KT had the original idea for the study and developed the intervention; UM, CL, and MB further contributed to the structure and tone of the intervention. MB undertook the computer programming of the intervention and performed data analyses. KT wrote the first draft of the manuscript to which MB contributed. All authors read and approved the final manuscript.

Conflicts of Interest

MB owns a private company that develops and distributes evidence-based lifestyle interventions to be used in health care settings. KT, UM, and CL declare no conflict of interest.

Multimedia Appendix 1
Table illustrating the themes of the program, number of text messages per week and examples of content.
[DOCX File , 16 KB - mhealth_v8i3e17208_app1.docx]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2885 KB - mhealth_v8i3e17208_app2.pdf]

References

1. World Health Organization. Mental health: strengthening our approach URL: http://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response [accessed 2019-11-26]
2. Keyes CLM, Annas J. Feeling good and functioning well: distinctive concepts in ancient philosophy and contemporary science. J Positive Psychol 2009 May;4(3):197-201. [doi: 10.1080/17439760902844228]
3. Taylor CT, Lyubomirsky S, Stein MB. Upregulating the positive affect system in anxiety and depression: Outcomes of a positive activity intervention. Depress Anxiety 2017 Mar 06;34(3):267-280. [doi: 10.1002/da.22593] [Medline: 28060463]
4. Keyes CL, Simoes EJ. To Flourish or Not: Positive Mental Health and All-Cause Mortality. Am J Public Health 2012 Nov;102(11):2164-2172. [doi: 10.2105/ajph.2012.300918]
5. Bolier L, Haverman M, Westerhof GJ, Riper H, Smit F, Bohlmeijer E. Positive psychology interventions: a meta-analysis of randomized controlled studies. BMC Public Health 2013 Feb 08;13:119 [FREE Full text] [doi: 10.1186/1471-2458-13-119] [Medline: 23390882]
6. Nikrahan GR, Laferton JA, Asgari K, Kalantari M, Abedi MR, Etesampour A, et al. Effects of Positive Psychology Interventions on Risk Biomarkers in Coronary Patients: A Randomized, Wait-List Controlled Pilot Trial. Psychosomatics 2016 Jul;57(4):359-368 [FREE Full text] [doi: 10.1016/j.psym.2016.02.007] [Medline: 27129358]
7. Grant F, Guille C, Sen S. Well-being and the risk of depression under stress. PLoS One 2013 Jul 1;8(7):e67395 [FREE Full text] [doi: 10.1371/journal.pone.0067395] [Medline: 23840872]
8. Keyes CLM, Dhingra SS, Simoes EJ. Change in Level of Positive Mental Health as a Predictor of Future Risk of Mental Illness. Am J Public Health 2010 Dec;100(12):2366-2371. [doi: 10.2105/ajph.2010.192245]
9. Schotanus-Dijkstra M, Ten Have M, Lamers SMA, de Graaf R, Bohlmeijer ET. The longitudinal relationship between flourishing mental health and incident mood, anxiety and substance use disorders. Eur J Public Health 2017 Jun 01;27(3):563-568. [doi: 10.1093/eurpub/ckw202] [Medline: 27818372]
10. Wood AM, Joseph S. The absence of positive psychological (eudemonic) well-being as a risk factor for depression: a ten year cohort study. J Affect Disord 2010 May;122(3):213-217. [doi: 10.1016/j.jad.2009.06.032] [Medline: 19706357]
11. Chida Y, Steptoe A. Positive Psychological Well-Being and Mortality: A Quantitative Review of Prospective Observational Studies. Psychosomatic Med 2008;70(7):741-756. [doi: 10.1097/psy.0b013e31818105ba]
12. Forsman AK, Wahlbeck K, Aarø LE, Alonso J, Barry MM, Brunn M, ROAMER Consortium. Research priorities for public mental health in Europe: recommendations of the ROAMER project. Eur J Public Health 2015 Apr 11;25(2):249-254. [doi: 10.1093/eurpub/cku232] [Medline: 25678606]

13. Keyes CLM. Promoting and protecting mental health as flourishing: a complementary strategy for improving national mental health. Am Psychol 2007;62(2):95-108. [doi: 10.1037/0003-066X.62.2.95] [Medline: 17324035]
14. Hone L, Jarden A, Schofield G. An evaluation of positive psychology intervention effectiveness trials using the re-aim framework: A practice-friendly review. J Positive Psychol 2014 Oct 13;10(4):303-322. [doi: 10.1080/17439760.2014.965267]
15. Sin NL, Lyubomirsky S. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis. J Clin Psychol 2009 May;65(5):467-487. [doi: 10.1002/jclp.20593] [Medline: 19301241]
16. Westerhof GJ, Keyes CLM. Mental Illness and Mental Health: The Two Continua Model Across the Lifespan. J Adult Dev 2010 Jun 26;17(2):110-119 [FREE Full text] [doi: 10.1007/s10804-009-9082-y] [Medline: 20502508]
17. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015 May 19;6(1):27882 [FREE Full text] [doi: 10.3402/ejpt.v6.27882] [Medline: 25994025]
18. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014 Aug 05;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
19. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
20. Versluis A, Verkuil B, Spinhoven P, van der Ploeg MM, Brosschot JF. Changing Mental Health and Positive Psychological Well-Being Using Ecological Momentary Interventions: A Systematic Review and Meta-analysis. J Med Internet Res 2016 Jun 27;18(6):e152 [FREE Full text] [doi: 10.2196/jmir.5642] [Medline: 27349305]
21. Karasouli E, Adams A. Assessing the Evidence for e-Resources for Mental Health Self-Management: A Systematic Literature Review. JMIR Ment Health 2014 Dec;1(1):e3 [FREE Full text] [doi: 10.2196/mental.3708] [Medline: 26543903]
22. Thomas K, Bendtsen M. Mental Health Promotion Among University Students Using Text Messaging: Protocol for a Randomized Controlled Trial of a Mobile Phone-Based Intervention. JMIR Res Protoc 2019 Aug 15;8(8):e12396 [FREE Full text] [doi: 10.2196/12396] [Medline: 31418426]
23. Keys CLM. Brief description of the mental health continuum short form (MHC-SF). 2009. URL: https://www.aacu.org/sites/default/files/MHC-SFEnglish.pdf [accessed 2019-11-26]
24. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
25. Lyubomirsky S, Layous K. How Do Simple Positive Activities Increase Well-Being? Curr Dir Psychol Sci 2013 Feb 22;22(1):57-62. [doi: 10.1177/0963721412469809]
26. SPIRIT Statement.: SPIRIT URL: https://www.spirit-statement.org/ [accessed 2020-02-26]
27. Conley CS, Durlak JA, Shapiro JB, Kirsch AC, Zahniser E. A Meta-Analysis of the Impact of Universal and Indicated Preventive Technology-Delivered Interventions for Higher Education Students. Prev Sci 2016 Dec;17(6):659-678. [doi: 10.1007/s11121-016-0662-3] [Medline: 27225631]
28. Weiss LA, Westerhof GJ, Bohlmeijer ET. Can We Increase Psychological Well-Being? The Effects of Interventions on Psychological Well-Being: A Meta-Analysis of Randomized Controlled Trials. PLoS One 2016 Jun 21;11(6):e0158092 [FREE Full text] [doi: 10.1371/journal.pone.0158092] [Medline: 27328124]
29. Schotanus-Dijkstra M, Drossaert CH, Pieterse ME, Boon B, Walburg JA, Bohlmeijer ET. An early intervention to promote well-being and flourishing and reduce anxiety and depression: A randomized controlled trial. Internet Interv 2017 Sep;9:15-24 [FREE Full text] [doi: 10.1016/j.invent.2017.04.002] [Medline: 30135833]
30. Müssener U, Bendtsen M, Karlsson N, White IR, McCambridge J, Bendtsen P. Effectiveness of Short Message Service Text-Based Smoking Cessation Intervention Among University Students: A Randomized Clinical Trial. JAMA Intern Med 2016 Mar 01;176(3):321-328 [FREE Full text] [doi: 10.1001/jamainternmed.2015.8260] [Medline: 26903176]
31. Thomas K, Müssener U, Linderoth C, Karlsson N, Bendtsen P, Bendtsen M. Effectiveness of a Text Messaging-Based Intervention Targeting Alcohol Consumption Among University Students: Randomized Controlled Trial. JMIR Mhealth Uhealth 2018 Jun 25;6(6):e146 [FREE Full text] [doi: 10.2196/mhealth.9642] [Medline: 29941417]
32. Huppert F. Psychological Well-being: Evidence regarding its causes and consequences. Appl Psychol Heal Well-being 2009;1(2):137-164. [doi: 10.1111/j.1758-0854.2009.01008.x]

Abbreviations

HADS: Hospital Anxiety Depression Scale
IRR: incidence rate ratio
MCAR: missing completely at random
MHC-SF: Mental Health Continuum Short Form
mHealth: mobile health
PPI: positive psychology intervention

Edited by G Eysenbach; submitted 26.11.19; peer-reviewed by T Hendriks, F Fries; comments to author 24.12.19; revised version received 31.01.20; accepted 07.02.20; published 20.03.20.

Please cite as:
Bendtsen M, Müssener U, Linderoth C, Thomas K
A Mobile Health Intervention for Mental Health Promotion Among University Students: Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(3):e17208
URL: http://mhealth.jmir.org/2020/3/e17208/
doi:10.2196/17208
PMID:32196462

Original Paper

Promoting Healthy Eating Habits for College Students Through Creating Dietary Diaries via a Smartphone App and Social Media Interaction: Online Survey Study

Masako Watanabe-Ito[1], MSN; Emiko Kishi[1], PhD; Yoko Shimizu[2], PhD

[1]Department of Public Health Nursing, School of Nursing, Toho University, Ota, Tokyo, Japan

[2]Department of Community Nursing, School of Nursing, Tokyo Women's Medical University, Shinjuku, Tokyo, Japan

Corresponding Author:
Masako Watanabe-Ito, MSN
Department of Public Health Nursing
School of Nursing
Toho University
4-16-20 Omori-Nishi
Ota, Tokyo, 143-0015
Japan
Phone: 81 3 3762 9881 ext 418
Fax: 81 3 3766 3914
Email: masako.watanabe@med.toho-u.ac.jp

Abstract

Background: Youth in developed countries face the contradictory health problems of obesity and an excessive desire for weight loss. Developing a better health attitude for college students is essential as this period of life establishes future lifestyle and habits. Online interaction on social media can help to improve eating habits by creating dietary diaries through a smartphone app; however, the effects of such interactions for college students have not been examined to date.

Objective: The aim of this study was to evaluate the potential effectiveness of social media interactions with the use of dietary diaries on a smartphone app to motivate college students in raising self-awareness of their eating habits.

Methods: Forty-two college students in the greater Tokyo area of Japan participated in the study by creating dietary diaries online through a smartphone app and then followed/interacted with each other using social media for 7 consecutive days in September to November 2017. Online surveys were administered at baseline, immediately after creating the dietary diaries, and at 1-month follow up. Participants rated their degree of interest and self-evaluation of eating habits using 7-point scales, and answered multiple choice questions related to their thoughts in choosing meals/drinks among 10 topics. Free descriptions about their overall experience throughout the project were also collected in the follow-up survey.

Results: Data from 38 participants who completed all processes were analyzed. Over time, the mean score for degree of interest in eating habits increased from 4.6 to 6.2 (P<.001), while the self-evaluation score decreased from 4.5 to 3.6 (P<.001); these significant differences remained after 1 month (5.3, P=.002; 4.1, P=0.04, respectively). A weak negative correlation (P=.009) was observed between scores for degree of interest and self-evaluation. Participants with lower scores for degree of interest at baseline tended to increase their interest level by more than 2 points above the average (P<.001). Participants gradually thought more about their eating habits from various perspectives when choosing a meal/drink, particularly with respect to maintaining well-balanced diets and introducing diverse ingredients. Participants evaluated their experiences as interesting/fun and reported familiarity with using the smartphone app and social media as the preferred method to keep track of their eating. All participants welcomed communication with fellow participants on social media and motivated each other, in addition to monitoring their eating habits through online dietary diaries. Some participants experienced difficulty, especially when they were busy or faced a lack of internet access.

Conclusions: Through interactions on social media, college students experienced encouragement and developed an interest and critical thinking with respect to their eating habits. This approach, which embraces peer education and peer support with social media, holds promise for the future of youth health promotion. Further examination will be needed to explore how to sustain this level of heightened awareness.

(JMIR Mhealth Uhealth 2020;8(3):e17613) doi:10.2196/17613

KEYWORDS

health promotion; college students; eating habits; social media; smartphone app

Introduction

In the last three decades, lifestyle-related health problems among youth in developed countries have become increasingly complicated given a simultaneous rise in the incidence of obesity and diabetes [1-3] with an excessive desire to lose weight by adopting unbalanced diets [4]. Among youth, college students are a particularly harder group to reach owing to their busy lives taken up by newly available activities associated with college life. Previous studies reported that although college students have adequate nutritional knowledge, their eating behavior is not necessarily healthy because they cannot recognize direct links between eating habits and health [5-8]. Therefore, it is essential to promote healthier eating habits among college students because lifestyles are established during this critical period, which have a significant impact on their future health.

Health education for college students requires new approaches that view young people as managers of their own eating habits rather than as recipients of health information. In Japan, 98.7% of people in their 20s use the internet, 88.7% own smartphones, and 78.5% use social media platforms such as Instagram and Twitter to communicate with each other and obtain information [9]. The World Health Organization also supports the potential of mobile health, which uses mobile and wireless technologies to support the achievement of health objectives, especially for motivating young people to acquire healthy behaviors [10]. Over the past 15 years, many researchers have adopted smartphone apps for health care such as for supporting behavioral management required during mental health care [11] and for self-monitoring in managing long-term conditions [12].

Since the 1980s, paper-based approaches such as food frequency questionnaires [13-15] and single or multiple daily recalls [16-18] have been conventionally applied for dietary assessment. Although these are cost-effective methods, some researchers noted that they are time consuming and rely on participants' memory and literacy, which can lead to higher rates of underreporting [19,20]. Since 1995, the use of innovative technologies has been shown to improve dietary assessment in various research settings. Research related to the use of dietary diaries with computer-based technologies has been conducted in both personal and interactive situations, demonstrating that data can be collected at a time that is convenient for participants [21,22]. Personal digital assistant technologies provide dietary diaries with a portion size measurement aid to enable participants to easily record their food [23,24], and mobile phone/smartphone-based technologies further enrich the data with the addition of digital photos and voice recording, as well as allowing easier registration regardless of location and time [25-27]. Illner et al [28] conducted a systematic review of the innovative technologies available for dietary diaries, demonstrating that dietary diaries that utilize technology have the potential to be more cost- and time-effective, and utilize less laborious means of data correction.

Public health research has expanded in recent years to explore methods that best promote a healthy diet and the adoption of information and communication technology, including a randomized controlled trial on the typical health specialist-patient relationship [29]. Research to promote college students' eating habits demonstrated that interventions employing information and communication technology can be effective [30-35], and some researchers adopted smartphones and personal digital assistants as assessment tools [36-38]. However, to date, few studies have examined the effects of online peer communication among participants monitoring their eating habits. Watanabe et al [39] investigated how college students interacted with each other through dietary diaries via an internet weblog that was accessed on flip-style phones; this approach was sufficiently familiar to enable participants to discover new challenges in their eating habits [39]. Turner-Mcgrievy and Tate [40] reported the effect of a weight loss program among adults using social media via smartphones. However, the effects of online interaction on social media by creating dietary diaries through a smartphone app to improve college students' eating habits have not yet been examined.

Our research explores how interactions through social media and creating dietary diaries with a smartphone app motivate college students to raise self-awareness of their eating habits in an effort to develop effective health education approaches for youth. In this study, we investigated (1) how college students change their interest levels and critical viewpoints toward their eating habits; (2) changes in various viewpoints with regard to their decision-making process when eating; and (3) their experiences from interactions on social media when creating dietary diaries via a smartphone app.

Methods

Research Design and Participants

This was a before-after study design conducted from September to November 2017 including 42 college students in the greater Tokyo area of Japan who were recruited through bulletin board posters at 5 cooperating universities. Any students at the cooperating universities under 25 years of age were eligible for inclusion in the study, regardless of gender or living situation. Following completion of all research processes, the participants received 2,000 JPY on prepaid cards that could be used at domestic convenience stores.

Procedure

Overall Design and Grouping

The participants were randomly divided into groups of 3 people and were asked to (1) create dietary diaries through the smartphone app and interact with/follow each other through social media; and (2) answer online surveys at baseline, immediately after creating the dietary diaries, and at 1-month follow up.

Creating Dietary Diaries

Participants recorded all of the food and drinks they consumed in their online diaries, including photos, text, and commentary space, during a 7-consecutive day period using a smartphone app. In the diaries, the app allowed them to register what they ate by choosing from a preregistered menu, products, or ingredients. Participants also wrote about their thoughts while eating or drinking. For example, they recorded the type of attention paid to choosing their meal or how they cooked the meal. Participants followed the diaries of their fellow group participants at least once a day, read blog-style diaries, and communicated with each other using the social media function in the commentary spaces of the diaries.

Online Surveys

Online surveys were administered three times: at baseline, immediately after the intervention, and at 1-month follow up. Participants were asked to respond to items on: (1) degree of interest in their eating habits, which was rated on a scale of 1 (“not at all”) to 7 (“very much”); (2) self-evaluation of their eating habits on a scale of 1 (“very bad”) to 7 (“very good”); and (3) multiple choice responses to the types of topics they considered when choosing food/drink among 10 topics based on a national survey of health and nutrition in Japan [41]. Data on basic personal characteristics and lifestyle were also collected at baseline. The final follow-up questionnaire included a free-response item asking for their overall experiences through participation in the project.

Research Settings and Smartphone App

Based on the pretest and our previous research [39], we set a 7-day intervention period so that participants had sufficient time to complete the required tasks, which included both keeping diaries and online surveys, and to ensure that their diaries reflected their eating habits depending on activities on both weekdays and weekends. We determined the number of participants in each group to allow for browsing participants’ diaries without unreasonable effort and to facilitate interaction with each other. We set the number of groups at 14 so that we could feasibly monitor the diaries and preempt any trouble between participants or unexpected disclosures of privacy. The smartphone app asken (asken Inc, Shinjuku, Tokyo, Japan) was used as the medium through which the participants created their dietary diaries and communicated with each other. The asken app is one of the most popular apps for diet management and nutrition improvement, with over 3.5 million users in Japan as of January 2020 [42]. We further selected the asken app for this research because it allows users to create diaries within a social networking system, thereby eliminating the need to send private messages, which safeguarded the participants’ privacy and security (Figure 1).

Figure 1. Details of the asken app.

Data Analysis

All 42 participants completed the diaries, but 4 were excluded from the analysis because they did not complete the 1-month follow-up online survey.

Quantitative data were analyzed using IBM SPSS Statistics, version 23 (SPSS Inc, Chicago, IL, USA). Changes in points for degree of interest, self-evaluation about their eating habits, and the number of topics that participants considered during their decision making were analyzed at three time points: baseline, immediately after creating dietary diaries, and at the 1-month follow up. First, Friedman tests were applied to determine whether any of the differences between the medians at the three time points were statistically significant. Since the *P* values were less than the significance threshold ($P<.01$), Wilcoxon signed-rank tests were used to examine the results between two time points compared with baseline, and Bonferroni correction was applied to the *P* values for multiple comparisons. To evaluate the degree of coherence between factors of participants who had a greater change in their scores for degree of interest than the average immediately after the intervention, Chi square tests were performed to compare basic characteristics and scores for degree of interest at baseline.

The content of free descriptions was analyzed using the qualitative content analysis method suggested by Graneheim and Lundman [43]. One author (MW) coded the content and discussed the results with other researchers until agreement was reached. The codes were then assigned to suitable categories and subcategories as agreed upon through discussion among researchers.

Ethical Considerations

This study was approved by the Ethical Review Board at Tokyo Women's Medical University (approval no. 4055, August 8, 2017). Participants were informed in writing of the research purpose and methods, that participation was voluntary, and that all collected data would be used only for research purposes and would be kept confidential. Participants were instructed not to upload any personal information online and provided written informed consent to participate. The investigators employed the latest security software to prevent data breaches, and all computers and documents were stored in secure areas.

Results

Participant Characteristics

The basic characteristics of the participants are summarized in Table 1.

All 38 participants were female, ranging in age from 19 to 22 years. Two participants were nutrition majors, and the remaining 36 majored in other subjects. The majority of participants had a self-reported body mass index in the normal range, whereas about 20% were underweight (mean 20.1, range 17.2-24.1) and none was overweight. The majority of participants lived with their families, followed by living alone and in college dormitories. All participants who lived with their families indicated that their parents cooked their meals, and just over half indicated that they cooked for themselves. The large majority of participants rated themselves to be in "excellent" or "good" health, while 3 participants rated their health as "not so good"; none rated their health as "poor."

Table 1. Basic characteristics of participants (N=38).

Characteristic	n (%)
Gender	
Male	0 (38)
Female	38 (100)
Age (years)	
19	8 (21)
20	7 (18)
21	13 (34)
22	10 (26)
Self-reported body mass index	
<18.5	8 (21)
18.5-25	30 (79)
>25	0 (0)
Living arrangement	
With family	23 (61)
Alone	11 (29)
In dormitory	4 (11)
People who cook daily meals (multiple choice)	
Family, especially parents	23 (61)
Themselves	20 (53)
Cafeteria	5 (13)
Eating out	8 (21)
Self-perceived health condition	
Excellent	12 (32)
Good	23 (61)
Not so good	3 (8)
Poor	0 (0)

Changes in Scores for Degree of Interest and Self-Evaluation of Eating Habits

Participants' scores for degree of interest in eating habits increased significantly (P<.001) while the scores for self-evaluation of eating habits decreased significantly (P<.001) throughout the process of creating dietary diaries. These differences remained significant at the 1-month follow up compared with the respective baseline values (Table 2).

The participants' interest level about their eating increased from baseline immediately after creating their dietary diaries, and then decreased slightly at the 1-month follow up. The average baseline score for self-evaluation of eating habits decreased immediately after creating dietary diaries and increased slightly at the 1-month follow up. A statistically significant but weak negative correlation was observed between scores of interest level and eating habits in self-evaluation (r=0.221, P=.009).

Regarding the factors that affected participants who changed their scores for degree of interest by more than the average, a smaller score at baseline (1-2) had a significant effect. However, there were no significant effects according to living arrangement, self-reported body mass index, cooking status, and self-perceived health condition (Table 3).

Table 2. Changes of participants' awareness associated with their eating habits through the intervention (N=38).

Item	Average	Median (range)	*P* value (Friedman test)	*P* value (Wilcoxan signed-rank test)[a]
Interest level in eating habits[b]			<.001	
Baseline	4.6	6 (2-5)		—[c]
Immediately after the project	6.2	6 (5-7)		<.001
One-month follow up	5.3	5 (2-7)		.002
Self-evaluation of eating habits[d]			.001	
Baseline	4.5	5 (2-6)		—
Immediately after the project	3.6	4 (1-7)		<.001
One-month follow up	4.1	4 (1-6)		.04
Number of topics considered when choosing a meal/drink[e]			<.001	
Baseline	3.7	2 (0-7)		—
Immediately after the project	6.4	5 (3-9)		<.001
One-month follow up	5.5	4 (2-8)		.05

[a]Compared with the baseline score after Bonferroni correction.
[b]Scale from 1 ("very bad") to 7 ("very good").
[c]Not applicable.
[d]Scale from 1 ("not at all") to 7 ("very much").
[e]Multiple choice answers.

Table 3. Factors affecting participants with substantial changes in their scores of interest level through the intervention (N=38).

Factor	Difference in points between baseline and after intervention		Chi-square	*P* value[a]
	>2, n (%)	≤2, n (%)		
Baseline score of interest level			11.17	.001
Stronger interest (5-6 at baseline)	2 (5)	14 (37)		
Weaker interest (1-4 at baseline)	16 (42)	6 (16)		
Living arrangement			0.07	.79
Alone/in dormitories	8 (21)	7 (18)		
With family	10 (26)	13 (34)		
Self-monitored body mass index			0.05	.82
<18.5	4 (11)	4 (11)		
≥18.5	14 (37)	16 (42)		
Responsible for cooking daily meals			0.01	.94
Themselves	8 (21)	12 (32)		
Others (family, cafeteria, eating out)	6 (16)	12 (32		
Self-perceived health condition			0.01	.94
Excellent/good	16 (42)	19 (50)		
Not so good/poor	2 (5)	1 (3)		

[a]After Yates' correction.

Changes in Dietary Topics Influencing Decision Making in Eating

The numbers of dietary topics participants thought about when they chose their meal/drink increased immediately after creating dietary diaries compared to that at baseline. The increase and the significant difference was maintained at the 1-month follow up (Table 2).

Considering the details of the topics more carefully (Table 4), at baseline, participants thought about "quantity of food consumed," "eating a variety of foods/ingredients," "whether

to drink alcohol or not," and "when to eat." After creating dietary diaries, the number of participants who chose "the nutrient balance of meals," "whether to eat breakfast or not," and "eating a variety of foods/ingredients" increased substantially. Eight participants chose "nothing in particular" at baseline, but no participant chose this response after creating dietary diaries.

Table 4. Categories considered during decision making about food and drink based on multiple choice responses (N=38).

Category	Baseline, n (%)	Immediately after the project, n (%)	One-month follow up, n (%)
Nothing in particular	8 (21)	0 (0)	0 (0)
Nutrient balance of meals	4 (11)	31 (82)	25 (66)
Eating a variety of foods/ingredients	14 (37)	23 (61)	24 (63)
Quantity of food consumed	17 (45)	15 (39)	14 (37)
How meals were cooked/processed	5 (13)	6 (16)	7 (18)
Time to eat	8 (21)	15 (39)	14 (37)
Eating breakfast or not	2 (5)	12 (32)	12 (32)
Eating out or not	4 (11)	8 (21)	6 (16)
Eating snacks/junk food or not	7 (18)	13 (34)	9 (24)
Drinking alcohol or not	8 (21)	10 (26)	10 (26)

Participants' Experience From the Project

Main Categories of Participant Experience

Participants' descriptions of their project experiences were sorted into 5 categories and 16 subcategories (Textbox 1), along with 52 lower-level codes from 223 meaning units.

Category A: "It was Fun/Interesting to Participate in the Project"

All participants responded that it was interesting to participate in the project, and their positive impressions about the project were included in Category A. They indicated that they enjoyed seeing what they and the other participants ate, taking photos, and reading comments and posts from other participants: "It was really fun to see what others ate, because we don't have a chance to see this in normal life. We are all college students, but live such different lives" (22-year old).

Category B: "I Learned From Participating in the Project"

Participants indicated that as a result of keeping their own diaries and observing others' diaries, they began to pay more attention to their eating habits such as how often they ate snacks, skipped breakfast, and ate/drank late at night: "I hadn't thought about eating, but through this project, I realized how terrible my eating habits are! How do I eat so many sweets in a day?" (21-year old).

Category C: "Participating Caused Some Difficulties"

Some participants wrote that participating was difficult because taking photos, writing comments in their diaries, and following others' diaries was time consuming and they sometimes forgot to record what they ate. They also wrote that they felt embarrassed to show their diaries to other participants when they did not eat well: "Even though it only took a little time, sometimes it was painful to record my meals in the diary, especially when I needed to do other things, such as study" (20-year old).

Category D: "Advantages/Disadvantages of Using the Smartphone App and Social Media"

Many participants mentioned that the smartphone app and social media were an advantage for health education because they were familiar with smartphone apps, and it matched their busy college lifestyles. They also found that using social media enabled them to communicate with each other, allowing them to exchange healthy eating tips such as how to include more vegetables in their meals. They also felt encouraged by other participants through comments and "good" posts to keep going.

Some participants wrote that participation was troublesome when they were without internet access or their smartphone batteries were low. One participant mentioned that it was a pity that they could not upload the project posts to the social media that they normally used to share content with friends. "To look back at our eating habits, using the smartphone app is a good idea! It is very familiar for us because we are living with it. On social media, we learned with each other. They encouraged me to complete it." (22-year old). "The app needed some time to get used to. Hopefully, it would be better if the recording method becomes easier." (19-year old).

Textbox 1. Qualitative analysis of participants' experiences of communication on social media when creating dietary diaries.

Category A: "It was fun/interesting to participate in the project"

A-1: Fun/interesting to keep the records of what I ate.

A-2: Fun/interesting to take photos of what I ate.

A-3: Fun/interesting to see what other members ate.

A-4: Fun/interesting to see comments and stamps from other members.

Category B: "I learned from participating in the project"

B-1: It made me think more of my eating habits and physical activities.

B-2: It made me more conscious of my eating habits.

B-3: I observed and learned from other members.

Category C: "Participating caused some difficulties"

C-1: It took time to complete project tasks.

C-2: It was difficult to record in general.

C-3: The period of the project was too short.

Category D: "Advantages/disadvantages of using the smartphone app and social media"

D-1: Advantages of using the mobile phone app include that it's familiar, always with me, and easy to record.

D-2: Disadvantages of using the mobile phone app include that it required special techniques and used up batteries.

D-3: Requests to improve the app.

D-4: Would like to use personal social media.

Category E: "My eating habits were affected by This project"

E-1: Improved eating habits during this project.

E-2: Improved eating habits even after this project.

Category E: "My Eating Habits Were Affected by this Project"

Some participants wrote that they changed their eating habits as a result of participating and began to eat breakfast every day, choose well-balanced meals with fresh vegetables/fruits, and avoid too many snacks and midnight eating. "After the project, I am trying to eat well-balanced meals as much as possible with many kinds of ingredients not only carbohydrates such as rice balls and bread" (21-year old).

Discussion

Principal Findings

This study represents an early attempt to explain how interaction using social media in conjunction with dietary diaries on a smartphone app motivates college students to develop interest in their eating habits. This intervention involved a multiplex process; nevertheless, the results show that college students experienced encouragement and developed an interest in their eating habits through interaction on social media when creating dietary diaries on the smartphone app. This methodology has potential as an effective means for youth to have a chance to review their eating habits for promoting healthier lifestyles.

Comparison With Prior Work

Participant Characteristics

Participant selection introduced some bias, resulting in only female participants. Young women who were concerned about their eating habits were more likely to be drawn to this research and to agree to participate. This was further influenced by the recruitment method because the collaborating colleges had departments with more female students. A well-developed recruitment plan is needed to enroll similar numbers of male students in future studies. Moreover, participants' basic characteristics such as self-reported body mass index, living arrangement, who cooked their meals, and their self-perceived health conditions did not differ from the average status of college students in the central-east area around Tokyo [44]. Participants' self-evaluations indicated that their baseline eating habits were not necessarily ideal, which was consistent with survey results from the Kanto Regional Agricultural Administration Office of Japan Ministry of Agriculture, Forestry and Fisheries [45], which showed that 42.2% of college students who live away from their families do not eat enough vegetables

and 71.3% want to improve their eating habits, suggesting the need for a new approach to providing health education on nutrition to this age group.

Verifying the Effectiveness of Dietary Diaries and Communication via Social Media: Increased Consciousness of Eating Habits

Participants' awareness of their eating habits increased, which was maintained during the project, although decreased slightly at the 1-month follow up. Participants became more interested in and thought critically about their eating habits, especially about eating nutrient-balanced meals and a variety of foods/ingredients. A previous study that incorporated use of an iPad reported that using photos and texts to record eating was an effective way to increase awareness of food intake [46]. Similarly, in our study, photos on diaries enabled participants to see at a glance how much they ate and how well-balanced the meal was, complementing their limited written descriptions. In addition, describing their thoughts when choosing a meal/drink in their diaries helped them to recognize what factors affect their decision making. Browsing and writing comments on other participants' diaries further provided an objective/new point of view to look back their own eating habits in comparative ways. In social learning theory, Bandura and Schunk [47] stated that people learn from each other through observation, imitation, and modeling. Our results support this theory of enhancing the peer-learning process but also suggest two challenges that warrant further examination: (1) how to sustain eating habit awareness/interest for a long period, and (2) how to motivate people to pay attention to the less visible factors such as how meals are processed.

Advantage of Using Social Media and Smartphone Apps to Promote Healthy Eating for College Students

The results showed that social media and smartphone apps have great potential for providing health education to youth because these tools are familiar to college students. Most participants evaluated the project as interesting and fun, and indicated that they discovered many tools for improving their eating habits. All participants also indicated that they felt motivated and encouraged by their fellow participants. This experience of peer support is important for health promotion, especially in this age group, as previously reported in the concept analysis of peer health support [48]. Wang and colleagues [49,50] developed a participatory action research strategy called "Photovoice" as an effective method for assessing the needs of social minorities. College students may similarly be seen as a group in need of help, given their lack of interest in and awareness of healthier eating habits. We suggest an approach such as "Photovoice-Online" using social media via smartphones to overcome this disadvantage, which requires participants to meet in person to discuss the problems with each other. Further research is needed to analyze the group dynamics of social media communication depending on specific group characteristics.

When social media is used in research, ethical considerations are an important concern. In particular, privacy and confidentiality safeguards are imperative when adopting social media for use in research. Holmberg et al [51] conducted a study about social media usage of adolescent patients with obesity, stating that social media could be a source for health inspiration, information, and support, but requires competencies. In this study, we adopted social media to encourage college students to be more conscious in their eating habits; however, in the setting of health education, more examination is needed to teach college students health literacy to use social media.

A few participants mentioned that it took time for them to record their meal/drink and communicate with others. Smartphone apps, which are developed and improved at a rapid pace, are already offering photo-based nutrient information, which will make the dietary diary recording process easier and more accurate. As desirable tools for college students, it would be preferable that more complementary smartphone apps including such high-level technologies would be available in the near future.

Limitations

This study has some limitations. As mentioned earlier, all participants in the study were young women, who are considered to be more conscious of their eating habits than young men. In addition, the study sample had a limited number of participants. Furthermore, as Deliens et al [7] noted, college students' eating habits are influenced by various factors such as social networks and physical/macro environments. As we were focused on the comparison with national Japanese results, we did not use the established questionnaires for international dietary assessment. Therefore, our results are not generalizable to larger populations. This intervention method adopted a multiplex process; thus, we could not analyze exactly which factors affected participants and to what extent. Finally, we did not examine participants' group dynamics, which may have also influenced the results.

Conclusion

This research explores how interactions through social media used in conjunction with smartphone apps of dietary diaries can motivate college students to develop an interest in healthier eating habits. Through interactions on social media when creating dietary diaries on a smartphone app, college students experienced encouragement and developed an interest in their eating. This methodology, which embraces peer education and peer support, holds promise for the future. A closer examination of group dynamics associated with participant interactions and longer-term experiments to develop sustainable motivation are needed to further advance this field of research.

Acknowledgments

We would like to express our appreciation to all the participants, and to asken Inc, which generously let us use their smartphone app and their images for our research. We are grateful to Dr Fumiko Miyaji and Professor Akiko Sasaki for their contributions.

This research was supported by grants from the Japan Society for Promoting Science Grants-in-Aid for Scientific Research, Grant No 26893283 (2014-2017) and 17K12549 (2017-2020).

Authors' Contributions

MW contributed to the conception and design of this study, performed the statistical/qualitative analysis, and drafted the manuscript. EK instructively engaged in data analysis with MW. EK and YS critically reviewed the manuscript and supervised the whole study process. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014 Aug 30;384(9945):766-781. [doi: 10.1016/S0140-6736(14)60460-8] [Medline: 24880830]
2. Singh GK, Siahpush M, Kogan MD. Rising social inequalities in US childhood obesity, 2003-2007. Ann Epidemiol 2010 Jan;20(1):40-52. [doi: 10.1016/j.annepidem.2009.09.008] [Medline: 20006275]
3. Centers for Disease Control Prevention (CDC). Physical activity levels among children aged 9-13 years--United States, 2002. MMWR Morb Mortal Wkly Rep 2003 Aug 22;52(33):785-788 [FREE Full text] [Medline: 12931076]
4. Eikey EV, Reddy MC, Booth KM, Kvasny L, Blair JL, Li V, et al. Desire to Be Underweight: Exploratory Study on a Weight Loss App Community and User Perceptions of the Impact on Disordered Eating Behaviors. JMIR Mhealth Uhealth 2017 Oct 12;5(10):e150 [FREE Full text] [doi: 10.2196/mhealth.6683] [Medline: 29025694]
5. Barnes SP, Brown KM, McDermott RJ, Bryant CA, Kromrey J. Perceived Parenting Style and the Eating Practices of College Freshmen. Am J Health Educ 2013 Jan 23;43(1):8-17. [doi: 10.1080/19325037.2012.10599213]
6. Japan Ministry of Health Labour and Welfare. 2019. The National health and nutrition survey in Japan, 2018 URL: https://www.mhlw.go.jp/stf/newpage_08789.html [accessed 2020-01-30]
7. Deliens T, Clarys P, De Bourdeaudhuij I, Deforche B. Determinants of eating behaviour in university students: a qualitative study using focus group discussions. BMC Public Health 2014 Jan 18;14:53 [FREE Full text] [doi: 10.1186/1471-2458-14-53] [Medline: 24438555]
8. Abraham S, Noriega BR, Shin JY. College students eating habits and knowledge of nutritional requirements. J Nutr Hum Health 2018;02(01):13-17 [FREE Full text] [doi: 10.35841/nutrition-human-health.2.1.13-17]
9. Japan Ministry of Internal Affairs and Communications. 2019. National ICT Survey in Japan URL: http://www.soumu.go.jp/johotsusintokei/statistics/data/190531_1.pdf [accessed 2020-01-30]
10. World Health Organization. WHO Executive Board 142nd session provisional agenda item 4.4. 2017 Nov 27. mHealth: Use of appropriate digital technologies for public health: report by the Director-General URL: https://apps.who.int/iris/bitstream/handle/10665/274134/B142_20-en.pdf?sequence=1&isAllowed=y [accessed 2020-01-30]
11. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: Integrating smartphone technology in behavioral healthcare. Prof Psychol Res Pract 2011 Dec;42(6):505-512. [doi: 10.1037/a0024485]
12. Whitehead L, Seaton P. The Effectiveness of Self-Management Mobile Phone and Tablet Apps in Long-term Condition Management: A Systematic Review. J Med Internet Res 2016;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
13. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol 2000 Feb 15;151(4):346-357 [FREE Full text] [doi: 10.1093/oxfordjournals.aje.a010213] [Medline: 10695593]
14. Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol 1997;26(Suppl 1):S6-S14. [doi: 10.1093/ije/26.suppl_1.s6] [Medline: 9126529]
15. Schatzkin A, Subar A, Thompson F, Harlan LC, Tangrea J, Hollenbeck AR, et al. Design and serendipity in establishing a large cohort with wide dietary intake distributions : the National Institutes of Health-American Association of Retired Persons Diet and Health Study. Am J Epidemiol 2001 Dec 15;154(12):1119-1125. [doi: 10.1093/aje/154.12.1119] [Medline: 11744517]
16. Dahm C, Keogh R, Spencer E, Greenwood DC, Key TJ, Fentiman IS, et al. Dietary fiber and colorectal cancer risk: a nested case-control study using food diaries. J Natl Cancer Inst 2010 May 05;102(9):614-626. [doi: 10.1093/jnci/djq092] [Medline: 20407088]
17. Schatzkin A, Kipnis V, Carroll R, Midthune D, Subar A, Bingham S, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. Int J Epidemiol 2003 Dec;32(6):1054-1062. [doi: 10.1093/ije/dyg264] [Medline: 14681273]

18. Bingham SA, Welch AA, McTaggart A, Mulligan AA, Runswick SA, Luben R, et al. Nutritional methods in the European Prospective Investigation of Cancer in Norfolk. Public Health Nutr 2001 Jun 02;4(3):847-858. [doi: 10.1079/phn2000102] [Medline: 11415493]
19. Prentice R, Mossavar-Rahmani Y, Huang Y, Van Horn L, Beresford S, Caan B, et al. Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment by using recovery biomarkers. Am J Epidemiol 2011 Sep 01;174(5):591-603 [FREE Full text] [doi: 10.1093/aje/kwr140] [Medline: 21765003]
20. Gibson RS, Charrondiere UR, Bell W. Measurement Errors in Dietary Assessment Using Self-Reported 24-Hour Recalls in Low-Income Countries and Strategies for Their Prevention. Adv Nutr 2017 Nov 15;8(6):980-991 [FREE Full text] [doi: 10.3945/an.117.016980] [Medline: 29141979]
21. Wong SS, Boushey CJ, Novotny R, Gustafson DR. Evaluation of a computerized food frequency questionnaire to estimate calcium intake of Asian, Hispanic, and non-Hispanic white youth. J Am Diet Assoc 2008 Mar;108(3):539-543. [doi: 10.1016/j.jada.2007.12.006] [Medline: 18313437]
22. Touvier M, Kesse-Guyot E, Méjean C, Pollet C, Malon A, Castetbon K, et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. Br J Nutr 2010 Nov 17;105(7):1055-1064. [doi: 10.1017/s0007114510004617]
23. Fowles ER, Gentry B. The feasibility of personal digital assistants (PDAs) to collect dietary intake data in low-income pregnant women. J Nutr Educ Behav 2008 Nov;40(6):374-377. [doi: 10.1016/j.jneb.2007.08.015] [Medline: 18984494]
24. Beasley J, Riley WT, Jean-Mary J. Accuracy of a PDA-based dietary assessment program. Nutrition 2005 Jun;21(6):672-677. [doi: 10.1016/j.nut.2004.11.006] [Medline: 15925290]
25. Wang D, Kogashiwa M, Kira S. Development of a new instrument for evaluating individuals' dietary intakes. J Am Diet Assoc 2006 Oct;106(10):1588-1593. [doi: 10.1016/j.jada.2006.07.004] [Medline: 17000191]
26. Pendergast FJ, Ridgers ND, Worsley A, McNaughton SA. Evaluation of a smartphone food diary application using objectively measured energy expenditure. Int J Behav Nutr Phys Act 2017 Dec 14;14(1):30 [FREE Full text] [doi: 10.1186/s12966-017-0488-9] [Medline: 28288657]
27. Ambrosini GL, Hurworth M, Giglia R, Trapp G, Strauss P. Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutr J 2018 Jan 09;17(1):5 [FREE Full text] [doi: 10.1186/s12937-018-0315-4] [Medline: 29316930]
28. Illner A, Freisling H, Boeing H, Huybrechts I, Crispim S, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 2012 Aug;41(4):1187-1203. [doi: 10.1093/ije/dys105] [Medline: 22933652]
29. Coughlin SS, Whitehead M, Sheats JQ, Mastromonico J, Hardy D, Smith SA. Smartphone Applications for Promoting Healthy Diet and Nutrition: A Literature Review. Jacobs J Food Nutr 2015;2(3):021 [FREE Full text] [Medline: 26819969]
30. Morris LM, Merrill RM. Use of the dietary analysis plus program in an introductory college nutrition class. J Nutr Educ Behav 2004;36(4):213-214. [doi: 10.1016/s1499-4046(06)60238-8] [Medline: 15544732]
31. Richards A, Kattelmann KK, Ren C. Motivating 18- to 24-year-olds to increase their fruit and vegetable consumption. J Am Diet Assoc 2006 Sep;106(9):1405-1411. [doi: 10.1016/j.jada.2006.06.005] [Medline: 16963345]
32. Franko L, Cousineau M, Trant M, Green TC, Rancourt D, Thompson D, et al. Motivation, self-efficacy, physical activity and nutrition in college students: randomized controlled trial of an internet-based education program. Prev Med 2008 Oct;47(4):369-377 [FREE Full text] [doi: 10.1016/j.ypmed.2008.06.013] [Medline: 18639581]
33. Clifford D, Anderson J, Auld G, Champ J. Good Grubbin': impact of a TV cooking show for college students living off campus. J Nutr Educ Behav 2009;41(3):194-200. [doi: 10.1016/j.jneb.2008.01.006] [Medline: 19411053]
34. Poddar KH, Hosig KW, Anderson ES, Nickols-Richardson SM, Duncan SE. Web-based nutrition education intervention improves self-efficacy and self-regulation related to increased dairy intake in college students. J Am Diet Assoc 2010 Nov;110(11):1723-1727. [doi: 10.1016/j.jada.2010.08.008] [Medline: 21034887]
35. Miller T, Chandler L, Mouttapa M. A Needs Assessment, Development, and Formative Evaluation of a Health Promotion Smartphone Application for College Students. Am J Health Educ 2015 Jul 06;46(4):207-215. [doi: 10.1080/19325037.2015.1044138]
36. West JH, Hall PC, Hanson CL, Barnes MD, Giraud-Carrier C, Barrett J. There's an app for that: content analysis of paid health and fitness apps. J Med Internet Res 2012;14(3):e72 [FREE Full text] [doi: 10.2196/jmir.1977] [Medline: 22584372]
37. Svensson A, Magnusson M, Larsson C. Overcoming Barriers: Adolescents' Experiences Using a Mobile Phone Dietary Assessment App. JMIR Mhealth Uhealth 2016 Jul 29;4(3):e92 [FREE Full text] [doi: 10.2196/mhealth.5700] [Medline: 27473462]
38. Brown ON, O'Connor LE, Savaiano D. Mobile MyPlate: a pilot study using text messaging to provide nutrition education and promote better dietary choices in college students. J Am Coll Health 2014;62(5):320-327. [doi: 10.1080/07448481.2014.899233] [Medline: 24654921]
39. Watanabe-Ito M, Miyaji F, Okamura J. Creating the dietary diary online: promoting healthier eating habits for youths. Journal of Okinawa Prefectural College of Nursing 2011;12:73-83 [FREE Full text]

40. Turner-McGrievy GM, Tate DF. Weight loss social support in 140 characters or less: use of an online social network in a remotely delivered weight loss intervention. Transl Behav Med 2013 Sep;3(3):287-294 [FREE Full text] [doi: 10.1007/s13142-012-0183-y] [Medline: 24073180]
41. Japan Ministry of Health Labour and Welfare. 2016. Overview of dietary reference intakes for Japanese in 2015 URL: https://www.mhlw.go.jp/file/06-Seisakujouhou-10900000-Kenkoukyoku/Overview.pdf [accessed 2020-03-20]
42. asken Inc. The asken App, make people's future healthier URL: https://www.asken.inc/eng/ [accessed 2020-03-20]
43. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
44. Kanto Regional Agricultural Administration Office of Japan Ministry of Agriculture, Forestry and Fisheries. 2014. Research of eating environment, eating habits interest for healthy eating about college/university students in Japanese Kanto region URL: http://www.maff.go.jp/kanto/press/syo_an/seikatsu/pdf/daigakuzentaiban.pdf [accessed 2020-03-20]
45. Japan Student Services Organization. 2017. Research of lives of college/university students in Japan URL: https://www.jasso.go.jp/about/statistics/gakusei_chosa/2018.html [accessed 2020-03-20]
46. Lister C, West JH, Richards R, Crookston B, Hall PC, Redelfs AH. Technology for health: A qualitative study on barriers to using the iPad for diet change. Health 2013;05(04):761-768. [doi: 10.4236/health.2013.54101]
47. Bandura A, Schunk DH. Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. J Person Soc Psychol 1981;41(3):586-598. [doi: 10.1037/0022-3514.41.3.586]
48. Dennis C. Peer support within a health care context: a concept analysis. Int J Nurs Stud 2003 Mar;40(3):321-332. [doi: 10.1016/S0020-7489(02)00092-5]
49. Wang CC, Yi WK, Tao ZW, Carovano K. Photovoice as a Participatory Health Promotion Strategy. Health Promot Int 1998 Jan 01;13(1):75-86. [doi: 10.1093/heapro/13.1.75]
50. Wang CC, Morrel-Samuels S, Hutchison PM, Bell L, Pestronk RM. Flint Photovoice: community building among youths, adults, and policymakers. Am J Public Health 2004 Jun;94(6):911-913. [doi: 10.2105/ajph.94.6.911] [Medline: 15249287]
51. Holmberg C, Berg C, Dahlgren J, Lissner L, Chaplin JE. Health literacy in a complex digital media landscape: Pediatric obesity patients' experiences with online weight, food, and health information. Health Informatics J 2018 Mar 01;25(4):1343-1357. [doi: 10.1177/1460458218759699] [Medline: 29499615]

Edited by G Eysenbach; submitted 29.12.19; peer-reviewed by JR Bautista, S King; comments to author 21.01.20; revised version received 22.02.20; accepted 11.03.20; published 31.03.20.

Please cite as:
Watanabe-Ito M, Kishi E, Shimizu Y
Promoting Healthy Eating Habits for College Students Through Creating Dietary Diaries via a Smartphone App and Social Media Interaction: Online Survey Study
JMIR Mhealth Uhealth 2020;8(3):e17613
URL: http://mhealth.jmir.org/2020/3/e17613/
doi:10.2196/17613
PMID:32229468

Original Paper

Evaluating the Feasibility and Acceptability of a Mobile Health–Based Female Community Health Volunteer Program for Hypertension Control in Rural Nepal: Cross-Sectional Study

Zhao Ni[1], PhD; Namratha Atluri[1], BSc; Ryan J Shaw[1], PhD; Jingru Tan[2], MPH; Kinza Khan[1]; Helena Merk[1], BSc; Yunfan Ge[2], MPH; Shrinkhala Shrestha[3], MPH; Abha Shrestha[3], MD; Lavanya Vasudevan[1], PhD; Biraj Karmacharya[3], PhD; Lijing L Yan[2], PhD, MPH

[1]Duke University, Durham, NC, United States
[2]Duke Kunshan University, Kunshan, China
[3]Kathmandu University, Kathmandu, Nepal

Corresponding Author:
Biraj Karmacharya, PhD
Kathmandu University
BP Highway, Panauti 45209
Kathmandu
Nepal
Phone: 977 11 490497
Email: birajmk@gmail.com

Related Article:

This is a corrected version. See correction statement: https://mhealth.jmir.org/2020/6/e19048/

Abstract

Background: Hypertension is a major modifiable risk factor for cardiovascular disease, the world's leading cause of death. The prevalence of hypertension is disproportionately higher in South Asian countries than in other regions of the world. Screening for hypertension in primary care settings remains a challenge in many South Asian countries, including Nepal. Nepal is located in the Himalayan Mountains region, posing significant geographical challenges for its rural citizens to access primary health care and service delivery. This barrier increases the costs and inconvenience for rural Nepalis to access hypertension screening and treatment. As a result, the prevalence of hypertension in Nepal tripled in the last 25 years to 22.4%-38.6%. Nepal's Ministry of Health and Population relies on female community health volunteers to link health centers and communities to provide basic health services. Over 50,000 of these volunteers in Nepal have received basic health care training and are assigned to take care of maternal and child health. Due to limited health care resources, adopting new methods to control hypertension is an urgent need in Nepal. Several recent studies in Nepal have recommended extending the role of female community health volunteers to include hypertension management through blood pressure monitoring and home-based education.

Objective: The goal of this study was to assess if a mobile health–based female community health volunteer approach of combining the traditional community health volunteer program with digital technologies would be feasible and acceptable in rural Nepal.

Methods: In this study, we recruited 17 female community health volunteers and extended their role from maternal and child health to hypertension management through screening blood pressures.

Results: All 17 female community health volunteers successfully measured 1113 rural Nepalis' blood pressures, identified 169 hypertensive patients, and collected health behaviors data of the 169 hypertensive patients. Among the 169 patients, 70% of them had a mobile phone, and 92% were interested in receiving health-related information via a mobile phone. Among those who were interested in receiving information via a mobile phone, 84% preferred voice calls, and 7% and 1% preferred texting and apps, respectively.

Conclusions: Results from this study indicate that a digital health intervention that leverages feature-phones combined with female community health volunteers may be an acceptable and pragmatic way to implement an evidence-based program to reduce hypertension in rural Nepal.

(JMIR Mhealth Uhealth 2020;8(3):e15419) doi:10.2196/15419

KEYWORDS

hypertension; female community health volunteers; mHealth

Introduction

Hypertension is a major modifiable risk factor for cardiovascular disease [1]. According to the Global Status Report on noncommunicable diseases, in the year 2010 alone, 9.4 million people died due to hypertension-related complications [2]. Compared to other regions of the world, the prevalence of hypertension is disproportionately higher in South Asian countries, such as Nepal. In Nepal in particular, the prevalence of hypertension tripled in the last 25 years to 22.4%-38.6% [3]. The mortality rate of hypertension has been steadily increasing in parallel, from 135.6 to 145.2 per 100,000 people from 1995 to 2015 [4]. Hypertension and related complications are considered major contributors to death and disability in Nepal [5]. Although hypertension can be effectively lowered using medication, there are several challenges to solving this issue in Nepal. First, there is inadequate screening for hypertension in primary care settings. Second, Nepal is mainly located in the Himalayan Mountains region, posing significant geographical challenges in access to primary health care and service delivery, which contributes to a high financial and human resource burden. Finally, among individuals identified as hypertensive, adherence to antihypertensive medication is consistently reported to be low, thereby increasing the risk of uncontrolled blood pressure and its complications [6].

Due to limited health care resources, adopting new methods to control hypertension is an urgent need in Nepal [7,8]. Nepal's Ministry of Health and Population relies on female community health volunteers to link health centers and communities to provide basic health services [9]. Over 50,000 female community health volunteers in Nepal have received basic health care training and are assigned to take care of maternal and child health [10]. Several studies in Nepal have recommended extending the role of these volunteers to include hypertension management through blood pressure monitoring and home-based education [9-11]. This recommendation is based on the fact that, in Nepal, there is a lack of linkage to primary care among patients with hypertension. Due to the high geographic barriers in the Himalayas, an affordable means to communicate and provide ready access to hypertension care delivery is greatly needed. Because mobile phones are widely used in Nepal, with a household subscription level of over 75% and steadily increasing [12], they may provide the needed communication platform. It is highly possible that female community health volunteers could use mobile phones to collect hypertension data and link patients to primary care. This suggests combining the traditional community health volunteer program with an evidence-based hypertension reduction program and digital technologies could be a pragmatic way to reduce hypertension.

The goal of this study was to assess if a mobile health (mHealth)–based female community health volunteer approach of combining the traditional volunteer program with digital technologies would be feasible and acceptable in rural Nepal.

Methods

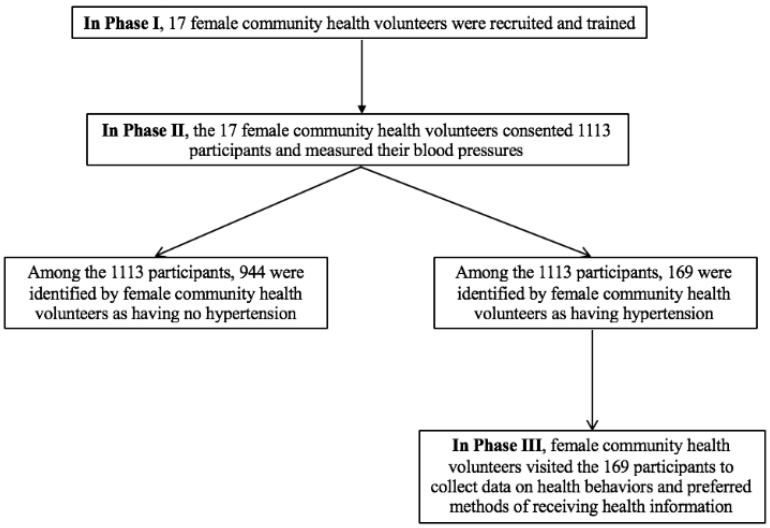
Study Design and Participants

This study was approved by the Duke University Health System Institutional Review Board (Pro00092469), and the ethics review committee in Nepal (Reg. no. 83/2018). This entire study was conducted in three phases (see Figure 1). In Phase I, we recruited 17 female community health volunteers from two rural communities, nine in Dhunkharka and eight in Panchkhal, Nepal, to learn how to accurately measure blood pressure. The government of Nepal granted permission to approach the 17 female community health volunteers for this study. All 17 female community health volunteers provided informed consent for the study. Each volunteer was responsible for the wards assigned by the Nepal Ministry of Health and Population. The two rural communities were located about 30 miles away from the capital city, Kathmandu. Two local Nepali graduate students who were fluent in English and Nepali served as translators for this study. Our team organized two equivalent training sessions taught by two licensed physicians.

Each training session included three subsessions: a 20-minute hypertension education session, a 10-minute question and answer session, and a 30-minute data-collection training session. Training content was developed specifically for this study by our study coordinator Jingru Tan and the two physicians by adapting content from the World Health Organization's Protocol titled "Prevention of Heart Attacks, Strokes, and Kidney Disease through Integrated Management of Diabetes and Hypertension" [13]. This protocol was part of the World Health Organization's package of essential noncommunicable disease interventions for primary health care in low-resource settings. In the training sessions, the female community health volunteers learned and practiced measuring, reading, and documenting blood pressure through an electronic blood pressure cuff. The volunteers also practiced measuring blood pressure on each other and received feedback from physicians. After the training sessions, our research team evaluated each community health volunteer's knowledge and practical skills on blood pressure assessment by assigning two team members to accompany each volunteer for half a day to observe their performance at measuring blood pressures. The two research team members evaluated the female community health volunteer's performance using the blood pressure–measuring guideline retrieved from the "Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure"

[14], which listed detailed information on how to measure blood pressure. After the observation, the two members provided feedback to each volunteer based on the blood pressure–measuring guideline and ensured that their skills for measuring blood pressure satisfied the guideline.

Figure 1. Study flow chart.



In Phase II, each female community health volunteer visited the rural residents in their respective wards and introduced this study. The inclusion criteria of participants were: (1) living in either of the two rural communities located in Dhunkharka and Panchkhal, Nepal for more than five years; (2) being able to speak Nepali or English; and (3) being 40 years old or older [13]. For those residents who provided written consent, the female community health volunteers measured their blood pressure twice on the right upper arm after 10 minutes of rest in a seated position at their homes, with feet on the floor and arm supported at heart level [14]. In total, 1159 residents satisfied the inclusion criteria, and 1113 of them were screened for hypertension, resulting in a recruitment rate of 96%. Of the 46 residents who refused to participate, 45 refused due to their busy schedules with farm work, and one refused without providing a specific reason. Participants with hypertension were selected for a survey in Phase III. In Phase III, female community health volunteers revisited the selected hypertensive participants, measured their blood pressure, and asked them to complete a 15-item survey. Each survey took about 15 minutes to complete.

Data Collection and Variables

In Phase II, volunteers collected participants' demographic variables, including gender, age, educational level, and health outcomes, including systolic blood pressure, diastolic blood pressure, and heart rate. The female community health volunteers visited each participant and measured blood pressure twice on the right upper arm after 10 minutes of rest in a seated position [15] by using an electronic blood pressure cuff (Omron HEM-7124; Omron Dalian Co. Ltd, China). The second measure was taken 5 minutes after the first measure. If the difference between the two measures was greater than 10 mmHg for systolic or diastolic blood pressure, a third measure was taken. The Omron blood pressure cuff has been widely used in Nepali clinical settings and is suitable for Nepali adults aged 18 years old or older. To ensure accuracy of measuring blood pressures, we used Omron blood pressure cuffs that have a bladder encircling at least 80 percent of the upper arm of an adult [14]. This Omron blood pressure cuff was a single size, but the selection of this cuff was based on its wide use in Nepali clinical settings and suitability for Nepali adults over 18 years of age [16,17]. The female community health volunteers can wrap the blood pressure cuff based on the upper arm circumference of a participant. In this study, we used the US Joint National Committee [14] definition of hypertension, in which hypertension is defined as average systolic blood pressure (SBP) equal to or greater than 140 mmHg, or an average diastolic blood pressure (DBP) equal to or greater than 90 mmHg [15,18]. The second visit in Phase III was to collect data on self-reported health behaviors, including smoking, alcohol consumption, medication adherence, hypertension awareness, phone use, and a preferred method of receiving health information (see Figure 1). The 15-item survey was adapted from the Morisky Green Levine Scale [19,20] and a 21-item survey querying from Johns Hopkins University School of Medicine [21].

Outcomes and Metrics

The primary outcome of this study was an assessment of the feasibility of extending the role of female community health volunteers to blood pressure monitoring. Feasibility was documented by the assignment completion rate, defined as the number of community health volunteers who completed all three phases. Secondary outcomes included systolic blood pressure, diastolic blood pressure, and preference of mHealth intervention. We also assessed the acceptability of this mHealth program, defined as the number of participants who consented to receive health-related information via cell phone, compared to the total number of participants approached in Phase III.

Statistical Analysis

All statistical analyses were performed using SAS 9.4 (SAS Institute, Cary, North Carolina, United States). Categorical variables included gender, educational level, marital status,

smoking, drinking consumption, and use of antihypertensive medications. Numerical variables included age, height, weight, systolic blood pressure, diastolic blood pressure, and heart rate. Descriptive statistics, frequency (n), and percentage (%) of female community health volunteers who completed all three phases, and participants who consented to receive health-related information via cell phone, were reported to detail the feasibility of extending the role of female community health volunteers to blood pressure monitoring, and the acceptability of this mHealth program among participants. The prevalence of hypertension, rate of medication adherence, percentage of mobile phone users, awareness of hypertension, and interest in receiving health-related information via mobile phones were also studied.

Results

The study sample included 1113 rural Nepalis with a mean age of 56.3 years old (SD 13.3), with a roughly equal distribution of male and female participants. Their ages ranged from 40-104 years old. Most participants (62%) had not received any formal education, and less than 6% had received higher education. Nearly one out of four participants (24%) had been diagnosed with hypertension by a doctor before the female community health volunteers measured their blood pressures; however, only 17% were taking medication to treat hypertension (Table 1). Therefore, among the participants who were diagnosed with hypertension by doctors, the percentage of them that were taking medications to treat hypertension was 69%.

Among the 1113 participants, 169 were identified by female community health volunteers to have hypertension in Phase II. The 169 participants included new hypertensive patients and hypertensive patients diagnosed by doctors who did not control their blood pressures within a normal range. The number of participants who were diagnosed and optimally treated was 189. Adding this number to the 169 hypertensive participants identified by the female community health volunteers, the total number of hypertensive participants was 358 (32%). The age range of the 169 participants was 40-104 years old. Their average systolic and diastolic blood pressures were 147 mmHg and 96 mmHg, respectively. Based on international hypertension guidelines [14,22], we categorized the 169 hypertensive participants into grade 1 (140≤SBP<160 mmHg and 90≤DBP<100 mmHg), and grade 2 (SBP≥160 mmHg and DBP≥100 mmHg). We found that 96 (57%) participants were in grade 1 and 73 (43%) were in grade 2. Among those 169 hypertensive rural Nepalis, 52% did not know that they had hypertension and 71% did not receive any treatment, such as medications to control their high blood pressure (Table 2). In terms of mobile phone usage, 70% of the hypertensive participants had a mobile phone, and 92% were interested in receiving health-related information via a mobile phone. Among those who were interested in receiving information via a mobile phone, 84% preferred voice calls, and 7% and 1% preferred texting and apps, respectively. All 17 female community health volunteers completed the three phases and returned their documented participants' blood pressures to us, suggesting the high feasibility of this approach. The acceptability was high, with 92% of the participants willing to receive health-related information via cell phone.

Table 1. Baseline characteristics of the sample in Phase II at enrollment.

Variable	All participants	Dhunkharka	Panchkhal	*P* value
Number of participants, n (%)	1113 (100)	594 (53.4)	519 (46.6)	
Gender[a], n (%)				**.49**
Male	520 (46.7)	283 (47.6)	237 (45.7)	
Female	592 (53.2)	310 (52.2)	282 (54.3)	
Education[b], n (%)				**.08**
No education	685 (61.5)	355 (59.8)	330 (63.6)	
Primary education	267 (24.0)	155 (26.1)	112 (21.6)	
Secondary education	92 (8.3)	41 (6.9)	51 (9.8)	
Higher education	64 (5.8)	38 (6.4)	26 (5.0)	
Had been diagnosed with hypertension by a doctor[c], n (%)				**.002**
No	842 (75.7)	428 (72.1)	414 (79.8)	
Yes	268 (24.1)	165 (27.8)	103 (19.8)	
Had been taking antihypertensive[d], n (%)				**<.001**
No	920 (82.7)	468 (78.8)	452 (87.1)	
Yes	185 (16.6)	125 (21.0)	60 (11.6)	
Age (Year), mean (SD)	56.3 (13.3)	55.5 (12.4)	57.3 (14.2)	.03

[a]The Gender information of one participant from Dhunkharka was missing.

[b]The Education information of five participants from Dhunkharka was missing.

[c]In Dhunkharka and Panchkhal, one and two participants' information on whether they had been diagnosed with hypertension by a doctor was missing, respectively.

[d]In Dhunkharka and Panchkhal, one and seven participants' information on whether they had been taking antihypertensives was missing, respectively.

Table 2. Baseline characteristics of the sample selected into Phase III.

Variable	All participants	Dhunkharka	Panchkhal	*P* Value
Number of participants, n (%)	169 (100)	68 (40.2)	101 (59.8)	__[a]
Gender, n (%)				**.49**
Male	99 (58.6)	42 (61.8)	57 (56.4)	
Female	70 (41.4)	26 (38.2)	44 (43.6)	
Education, n (%)				**.13**
No education	99 (58.6)	35 (51.5)	64 (63.4)	
Primary education	46 (27.2)	23 (33.8)	23 (22.8)	
Secondary education	10 (5.9)	3 (4.4)	7 (6.9)	
Higher education	11 (6.5)	7 (10.3)	4 (4.0)	
Had been diagnosed with hypertension by a doctor (received prescription of antihypertensive medications), n (%)				**.03**
No	88 (52.1)	42 (61.8)	46 (45.5)	
Yes	79 (46.7)	25 (36.8)	54 (53.5)	
Had been taking antihypertensive, n (%)				**.01**
No	120 (71.0)	56 (82.4)	64 (63.4)	
Yes	45 (26.6)	11 (16.2)	34 (33.7)	
Stopped taking medication when felt worse[b], n (%)				**.01**
No	41 (91.1)	8 (72.7)	33 (97.1)	
Yes	3 (6.7)	3 (27.3)	0 (0)	
Stopped taking medication when felt better[b], n (%)				**.05**
No	39 (86.7)	8 (72.7)	31 (91.2)	
Yes	5 (11.1)	3 (27.3)	2 (5.9)	
Perspective on whether hypertensive patients need to take antihypertensive daily, n (%)				**.01**
No	37 (21.9)	22 (32.4)	15 (14.9)	
Yes	114 (67.5)	39 (57.4)	75 (74.3)	
Had smoked before, n (%)				**.57**
No	81 (47.9)	31 (45.6)	50 (49.5)	
Yes	82 (48.5)	35 (51.5)	47 (46.5)	
Current smoker[c], n (%)				**.80**
No	30 (36.6)	14 (40.0)	16 (34.0)	
Yes	39 (47.6)	17 (48.6)	22 (46.8)	
Drinking alcohol, n (%)				**.03**
Never	105 (62.1)	38 (55.9)	67 (66.3)	
Drink during events	31 (18.3)	12 (17.6)	19 (18.8)	
1-2 times a month	6 (3.6)	3 (4.4)	3 (3.0)	
1-2 times a week	4 (2.4)	2 (2.9)	2 (2.0)	
Daily	15 (8.9)	12 (17.6)	3 (3.0)	
Had a cell phone, n (%)				**.64**
No	44 (26.0)	16 (23.5)	28 (27.7)	
Yes	119 (70.4)	48 (70.6)	71 (70.3)	
The phone is a smartphone[d], n (%)				**.39**
No	81 (68.1)	34 (70.8)	47 (66.2)	

Variable	All participants	Dhunkharka	Panchkhal	*P* Value
Yes	33 (27.7)	11 (22.9)	22 (31.0)	
Interested in receiving health-related information, n (%)				**.01**
No	7 (4.1)	6 (8.8)	1 (1.0)	
Yes	156 (92.3)	59 (86.8)	97 (96.0)	
Preferred methods,[e] n (%)				**.13**
Voice call	131 (84.0)	50 (84.7)	81 (83.5)	
Text message	11 (7.1)	2 (3.4)	9 (9.3)	
Apps	2 (1.3)	2 (3.4)	0 (0)	
Don't know	1 (0.6)	0 (0)	1 (1.0)	
Weight (kg), mean (SD)	60.9 (12.5)	58.4 (10.5)	62.8 (13.6)	.05
Height (cm), mean (SD)	159.7 (11.0)	157.6 (14.7)	160.7 (8.6)	.20
Age (years), mean (SD)	59.3 (13.3)	57.0 (10.7)	60.8 (14.7)	.07
BMI (kg/m^2), mean (SD)	24.6 (5.7)	24.3 (5.4)	24.8 (5.9)	.71

[a]Not applicable.

[b]Answers were from participants who had been taking antihypertensive.

[c]Answers were from participants who had smoked before.

[d]Answers were from participants who had a cell phone.

[e]Answers were from participants who had interest in receiving health-related information.

Discussion

Primary Findings

This study demonstrates high feasibility and acceptability of using female community health volunteers, an integral component of the Nepali health care infrastructure, in collecting blood pressure data. Although female community health volunteers in Nepal have mainly been used for maternal and child health interventions, our study shows that these volunteers can be trained to assist with other important health initiatives. In addition to quickly learning how to use electronic blood pressure cuffs and record and interpret blood pressure readings, the female community health volunteers were also keen on imparting their newly gained knowledge about hypertension to their community members. Globally, the role of community health workers (CHWs) has shifted from serving a specific disease or population, such as maternal and child health, to help solve community health problems that are the most emergent or need the most resources [23]. Hypertension is a major risk factor for cardiovascular disease, the world's leading cause of death. Bone et al [24] and Krieger et al [25] have extended the role of a CHW to include controlling high blood pressure and demonstrated that CHWs could successfully increase follow-up care for hypertensive patients. Our findings are aligned with the existing evidence. Shifting primary care duties of managing hypertension in low- and middle- income countries from physicians to nonphysician health care workers, such as CHWs, has been tested in studies and showed potential in reducing blood pressure [26]. For example, a high-quality randomized trial conducted in rural China and India [27] found that CHWs using an Android-powered app could reduce blood pressure and improve antihypertensive medication adherence among rural residents with cardiovascular disease. Studies conducted in Pakistan [28], Bangladesh, and Sri Lanka [29] also found the same result that CHW-led interventions can help reduce blood pressure.

The results from this study indicate that hypertension management remains grossly inadequate in rural Nepal. Among the 169 female community health volunteer–identified hypertensive participants, 52% were not aware of their hypertension and nearly 71% did not regularly take medications to control their hypertension. Overall, 22% of the participants were not aware that hypertensive patients need to take antihypertensive medications to manage high blood pressure. This data reveals how uninformed participants are about their blood pressure, how to manage it through medication adherence, and the risks involved with living with this chronic illness. Medication adherence can be jointly influenced by external factors such as health systems, health providers, and access to care, and intrinsic factors associated with patients. Intrinsic nonadherence is caused primarily by forgetfulness, misunderstanding of the medication regimens, and lack of communication with health workers [30,31]. While several medication adherence–increasing interventions exist [32], most are ill-suited for low-income countries lacking necessary health care infrastructure [6].

In this study's participant population, in addition to the financial disadvantage indicated by their reported income levels, participants also faced geographic isolation. We found through our study that participants' access to primary health care facilities is very limited, which can contribute to their lack of understanding of how to self-manage their health and the consequences of poor medication adherence. Due to this inconsistent relationship with the health care system, many rural

Nepalis in our study also express distrust of medical professionals and are less inclined to seek out health services on their own. However, the female community health volunteers are highly trusted throughout their communities and have close relationships with residents. This study has shown that the female community health volunteers can implement hypertension-based health initiatives; thus, future interventions to reduce hypertension in Nepal could leverage female community health volunteers for community-based blood pressure monitoring. In this study, the electronic blood pressure cuffs that we used were powered by batteries, which were affordable and accessible in rural Nepal. For example, a set of four regular batteries was 80 Nepalese Rs (US $0.70). Today, there are many types of blood pressure monitors available in the market. For instance, those with a USB connection can be charged by portable batteries and can provide users the guidance on acceptable blood pressure ranges. They can be adapted in rural Nepal by female community health volunteers too.

The collected data also revealed that 70% of participants owned a feature-phone that had text messaging abilities. Overall, 84% of the participants also preferred receiving health information through voice calls or voice messages on their phones Considering the huge cost barrier to improving health outcomes in rural Nepal, both from an individual patient perspective and a governmental perspective, these data suggest a new avenue to deliver care by utilizing mobile health technologies [33,34]. Since lack of awareness about health information and insufficient medication adherence were shown to be problematic for hypertensive patients in rural Nepal, the high prevalence of feature-phone use and patients' interest in getting health information via voice calls provide a strong rationale to support a feature phone-based software application. Integrating such an intervention with existing infrastructure, such as the female community health volunteer system, appears promising as the volunteers are already highly trusted and have a wide reach. For example, after visiting a hypertensive resident, the female community health volunteer may use phone calls to follow up on the resident's access to antihypertensive medications. When refilling medications is necessary, the female community health volunteers can bring medications to that resident in their next visit. Also, the cell phones may be used to facilitate female community health volunteers to monitor rural residents' blood pressures and thus increase early diagnosis of hypertension. For example, by following up on residents and asking about their previously tested blood pressures (tested in a clinic or hospital) via cell phones, the volunteers could find out if the residents' blood pressures had increased. If a resident's blood pressure is increasing, the female community health volunteers can visit that resident and measure his or her blood pressure to confirm whether the resident is hypertensive.

The technology considered in this feature phone–based intervention should refer to basic mobile phones, which will enable broader coverage in low-income countries such as Nepal. These devices lack smartphone capabilities and tend to have custom-designed software and user interfaces. Many studies have confirmed the effectiveness and affordability of text message reminders to improve adherence to medication intake and to encourage regular visit attendance to manage chronic diseases [35]. Also, Nepal's ministry of health has prioritized mHealth-led interventions by releasing "National e-health strategies," aiming to strengthen the application of information and communications technologies in support of health and health-related fields [36]. When assessed together, the aforementioned factors demonstrate an intervention to reduce hypertension that leverages feature-phones as the communication platform could possibly be a well-timed program for Nepal.

While this study contributes important knowledge to the use of mHealth in rural Nepal, it has several limitations. First, the female community health volunteers identified hypertensive participants by only a single visit. Given the high variability in blood pressure, ideally, the volunteers should visit participants twice to measure their blood pressure. Second, in this study, variables such as weight and drinking alcohol were not measured by female community health volunteers but were self-reported by participants. Also, a challenge in the management of chronic conditions, such as hypertension and diabetes, in Nepal is access to medications and a patient's ability to pay for them. In this study, we did not address this issue. Finally, our analyses were based on the 169 hypertensive patients identified by the female community health volunteers but did not include those diagnosed by doctors and optimally controlled their blood pressures within the normal range.

Conclusion

Given the large burden of hypertension in Nepal, adopting new methods to control hypertension has become an emergent need for rural Nepalese. Results from this study indicate that a mobile health intervention that leverages feature-phones combined with female community health volunteers is a new way to implement an evidence-based program to reduce hypertension in rural Nepal. More future research should be conducted to test the feasibility and acceptability of such programs.

Conflicts of Interest

None declared.

References

1. Cappuccio FP, Miller MA. Cardiovascular disease and hypertension in sub-Saharan Africa: burden, risk and interventions. Intern Emerg Med 2016 Apr;11(3):299-305 [FREE Full text] [doi: 10.1007/s11739-016-1423-9] [Medline: 27001886]

2. Mendis S, Armstrong T, Bettcher D, Branca F, Lauer J, Mace C, et al. World Health Organization. 2014. Global status report on noncommunicable diseases 2014 URL: https://apps.who.int/iris/bitstream/handle/10665/148114/9789241564854_eng.pdf;jsessionid=7723E6F66DD9FA4E4E07F30ED220DF24?sequence=1 [accessed 2018-12-11]
3. Mishra SR, Neupane D, Bhandari PM, Khanal V, Kallestrup P. Burgeoning burden of non-communicable diseases in Nepal: a scoping review. Global Health 2015 Jul 16;11:32 [FREE Full text] [doi: 10.1186/s12992-015-0119-7] [Medline: 26178459]
4. Forouzanfar MH, Liu P, Roth GA, Ng M, Biryukov S, Marczak L, et al. Global Burden of Hypertension and Systolic Blood Pressure of at Least 110 to 115 mm Hg, 1990-2015. JAMA 2017 Jan 10;317(2):165-182. [doi: 10.1001/jama.2016.19043] [Medline: 28097354]
5. Hasan M, Sutradhar I, Akter T, Das Gupta R, Joshi H, Haider MR, et al. Prevalence and determinants of hypertension among adult population in Nepal: Data from Nepal Demographic and Health Survey 2016. PLoS One 2018;13(5):e0198028 [FREE Full text] [doi: 10.1371/journal.pone.0198028] [Medline: 29852006]
6. Schroeder K, Fahey T, Ebrahim S. How can we improve adherence to blood pressure-lowering medication in ambulatory care? Systematic review of randomized controlled trials. Arch Intern Med 2004 Apr 12;164(7):722-732. [doi: 10.1001/archinte.164.7.722] [Medline: 15078641]
7. Karmacharya BM, Koju RP, LoGerfo JP, Chan KCG, Mokdad AH, Shrestha A, et al. Awareness, treatment and control of hypertension in Nepal: findings from the Dhulikhel Heart Study. Heart Asia 2017;9(1):1-8 [FREE Full text] [doi: 10.1136/heartasia-2016-010766] [Medline: 28123454]
8. Khanal MK, Dhungana RR, Bhandari P, Gurung Y, Paudel KN. Prevalence, associated factors, awareness, treatment, and control of hypertension: Findings from a cross sectional study conducted as a part of a community based intervention trial in Surkhet, Mid-western region of Nepal. PLoS One 2017;12(10):e0185806 [FREE Full text] [doi: 10.1371/journal.pone.0185806] [Medline: 28982159]
9. Neupane D, Kallestrup P, McLachlan CS, Perry H. Community health workers for non-communicable diseases. Lancet Glob Health 2014 Oct;2(10):e567 [FREE Full text] [doi: 10.1016/S2214-109X(14)70303-1] [Medline: 25304630]
10. Neupane D, Mclachlan CS, Gautam R, Mishra SR, Thorlund M, Schlütter M, et al. Literacy and motivation for the prevention and control of hypertension among female community health volunteers: a qualitative study from Nepal. Glob Health Action 2015;8:28254 [FREE Full text] [doi: 10.3402/gha.v8.28254] [Medline: 26679406]
11. Neupane D, McLachlan CS, Mishra SR, Olsen MH, Perry HB, Karki A, et al. Effectiveness of a lifestyle intervention led by female community health volunteers versus usual care in blood pressure reduction (COBIN): an open-label, cluster-randomised trial. Lancet Glob Health 2018 Jan;6(1):e66-e73 [FREE Full text] [doi: 10.1016/S2214-109X(17)30411-4] [Medline: 29241617]
12. Wilson R, Zu Erbach-Schoenberg E, Albert M, Power D, Tudge S, Gonzalez M, et al. Rapid and Near Real-Time Assessments of Population Displacement Using Mobile Phone Data Following Disasters: The 2015 Nepal Earthquake. PLoS Curr 2016 Feb 24;8 [FREE Full text] [doi: 10.1371/currents.dis.d073fbece328e4c39087bc086d694b5c] [Medline: 26981327]
13. Chestnov O, Mendis S, Roglic G, Sepulveda C, Shao R. World Health Organization. 2013. Implementation tools: package of essential noncommunicable (PEN) disease interventions for primary health care in low-resource settings URL: https://apps.who.int/iris/bitstream/handle/10665/133525/9789241506557_eng.pdf?sequence=1 [accessed 2018-12-11]
14. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, et al. National Institutes of Health. 2004 Aug. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (Complete Report) URL: https://www.ncbi.nlm.nih.gov/books/NBK9630/pdf/Bookshelf_NBK9630.pdf [accessed 2018-12-30]
15. Lu J, Lu Y, Wang X, Li X, Linderman GC, Wu C, et al. Prevalence, awareness, treatment, and control of hypertension in China: data from 1·7 million adults in a population-based screening study (China PEACE Million Persons Project). Lancet 2017 Dec 09;390(10112):2549-2558. [doi: 10.1016/S0140-6736(17)32478-9] [Medline: 29102084]
16. Chahine MN, Topouchian J, Zelveian P, Hakobyan Z, Melkonyan A, Azaki A, et al. Validation of BP devices QardioArm in the general population and Omron M6 Comfort in type II diabetic patients according to the European Society of Hypertension International Protocol (ESH-IP). Med Devices (Auckl) 2018;11:11-20 [FREE Full text] [doi: 10.2147/MDER.S142126] [Medline: 29343992]
17. Aryal KK, Mehata S, Neupane S, Vaidya A, Dhimal M, Dhakal P, et al. The Burden and Determinants of Non Communicable Diseases Risk Factors in Nepal: Findings from a Nationwide STEPS Survey. PLoS One 2015;10(8):e0134834 [FREE Full text] [doi: 10.1371/journal.pone.0134834] [Medline: 26244512]
18. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, National Heart, Lung,Blood Institute Joint National Committee on Prevention, Detection, Evaluation,Treatment of High Blood Pressure, National High Blood Pressure Education Program Coordinating Committee. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. JAMA 2003 May 21;289(19):2560-2572. [doi: 10.1001/jama.289.19.2560] [Medline: 12748199]
19. Unni EJ, Farris KB. Development of a new scale to measure self-reported medication nonadherence. Res Social Adm Pharm 2015;11(3):e133-e143. [doi: 10.1016/j.sapharm.2009.06.005] [Medline: 21272524]
20. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986 Jan;24(1):67-74. [doi: 10.1097/00005650-198601000-00007] [Medline: 3945130]

21. Leite L, Buresh M, Rios N, Conley A, Flys T, Page KR. Cell phone utilization among foreign-born Latinos: a promising tool for dissemination of health and HIV information. J Immigr Minor Health 2014 Aug;16(4):661-669 [FREE Full text] [doi: 10.1007/s10903-013-9792-x] [Medline: 23440452]
22. Mancia G, Fagard R, Narkiewicz K, Redón J, Zanchetti A, Böhm M, Task Force Members. 2013 ESH/ESC Guidelines for the management of arterial hypertension: the Task Force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). J Hypertens 2013 Jul;31(7):1281-1357. [doi: 10.1097/01.hjh.0000431740.32696.cc] [Medline: 23817082]
23. Swider SM. Outcome effectiveness of community health workers: an integrative literature review. Public Health Nurs 2002;19(1):11-20. [doi: 10.1046/j.1525-1446.2002.19003.x] [Medline: 11841678]
24. Bone LR, Mamon J, Levine DM, Walrath JM, Nanda J, Gurley HT, et al. Emergency department detection and follow-up of high blood pressure: use and effectiveness of community health workers. Am J Emerg Med 1989 Jan;7(1):16-20. [doi: 10.1016/0735-6757(89)90077-6] [Medline: 2914043]
25. Krieger J, Collier C, Song L, Martin D. Linking community-based blood pressure measurement to clinical care: a randomized controlled trial of outreach and tracking by community health workers. Am J Public Health 1999 Jun;89(6):856-861. [doi: 10.2105/ajph.89.6.856] [Medline: 10358675]
26. Anand TN, Joseph LM, Geetha AV, Prabhakaran D, Jeemon P. Task sharing with non-physician health-care workers for management of blood pressure in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Glob Health 2019 Jun;7(6):e761-e771 [FREE Full text] [doi: 10.1016/S2214-109X(19)30077-4] [Medline: 31097278]
27. Tian M, Ajay VS, Dunzhu D, Hameed SS, Li X, Liu Z, et al. A Cluster-Randomized, Controlled Trial of a Simplified Multifaceted Management Program for Individuals at High Cardiovascular Risk (SimCard Trial) in Rural Tibet, China, and Haryana, India. Circulation 2015 Sep 01;132(9):815-824 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.015373] [Medline: 26187183]
28. Jafar TH, Hatcher J, Poulter N, Islam M, Hashmi S, Qadri Z, Hypertension Research Group. Community-based interventions to promote blood pressure control in a developing country: a cluster randomized trial. Ann Intern Med 2009 Nov 03;151(9):593-601. [doi: 10.7326/0003-4819-151-9-200911030-00004] [Medline: 19884620]
29. Jafar TH, Silva AD, Naheed A, Jehan I, Liang F, Assam PN, COBRA-BPS Study Group. Control of blood pressure and risk attenuation: a public health intervention in rural Bangladesh, Pakistan, and Sri Lanka: feasibility trial results. J Hypertens 2016 Sep;34(9):1872-1881. [doi: 10.1097/HJH.0000000000001014] [Medline: 27488552]
30. Lehane E, McCarthy G. Intentional and unintentional medication non-adherence: a comprehensive framework for clinical research and practice? A discussion paper. Int J Nurs Stud 2007 Nov;44(8):1468-1477. [doi: 10.1016/j.ijnurstu.2006.07.010] [Medline: 16973166]
31. Egan BM, Zhao Y, Axon RN. US trends in prevalence, awareness, treatment, and control of hypertension, 1988-2008. JAMA 2010 May 26;303(20):2043-2050. [doi: 10.1001/jama.2010.650] [Medline: 20501926]
32. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keepanasseril A, et al. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2014 Nov 20(11):CD000011. [doi: 10.1002/14651858.CD000011.pub4] [Medline: 25412402]
33. 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 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
34. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med 2013;10(2):e1001382 [FREE Full text] [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]
35. Yasmin F, Banu B, Zakir SM, Sauerborn R, Ali L, Souares A. Positive influence of short message service and voice call interventions on adherence and health outcomes in case of chronic disease care: a systematic review. BMC Med Inform Decis Mak 2016 Apr 22;16:46 [FREE Full text] [doi: 10.1186/s12911-016-0286-3] [Medline: 27106263]
36. Government of Nepal. 2017. National e-Health Strategy 2017 URL: https://www.mohp.gov.np/attachments/article/368/Nepal_e_health_strategy_Eng.pdf [accessed 2019-06-23]

Abbreviations

CHW: community health worker
DBP: diastolic blood pressure
mHealth: mobile health
SBP: systolic blood pressure

Edited by G Eysenbach; submitted 10.07.19; peer-reviewed by V Cornelissen, B Green; comments to author 01.09.19; revised version received 29.10.19; accepted 17.12.19; published 09.03.20.

Please cite as:
Ni Z, Atluri N, Shaw RJ, Tan J, Khan K, Merk H, Ge Y, Shrestha S, Shrestha A, Vasudevan L, Karmacharya B, Yan LL
Evaluating the Feasibility and Acceptability of a Mobile Health–Based Female Community Health Volunteer Program for Hypertension Control in Rural Nepal: Cross-Sectional Study
JMIR Mhealth Uhealth 2020;8(3):e15419
URL: http://mhealth.jmir.org/2020/3/e15419/
doi:10.2196/15419
PMID:32149712

Original Paper

Using Mobile Health Tools to Engage Rural Underserved Individuals in a Diabetes Education Program in South Texas: Feasibility Study

Zenong Yin[1], MA, PhD; Janna Lesser[2], RN, FAAN, PhD; Kristi A Paiva[2], MPH; Jose Zapata Jr[2], MSS; Andrea Moreno-Vasquez[2], MPH; Timothy J Grigsby[1], PhD; Stacy R Ryan-Pettes[3], PhD; Deborah Parra-Medina[4], MPH, PhD, FAAHB; Vanessa Estrada[1], MSc; Shiyu Li[1], MSc; Jing Wang[2], MPH, PhD, RN, FAAN

[1]Department of Public Health, College of Health, Community and Policy, University of Texas at San Antonio, San Antonio, TX, United States

[2]School of Nursing, The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

[3]Department of Psychology and Neuroscience, Baylor University, Waco, TX, United States

[4]Department of Mexican American and Latina/o Studies, The University of Texas at Austin, Austin, TX, United States

Corresponding Author:
Jing Wang, MPH, PhD, RN, FAAN
School of Nursing
The University of Texas Health Science Center at San Antonio
7703 Floyd Curl Drive
San Antonio, TX, 78229
United States
Phone: 1 210 450 8561
Email: wangj1@uthscsa.edu

Abstract

Background: Access to diabetes education and resources for diabetes self-management is limited in rural communities, despite higher rates of diabetes in rural populations compared with urban populations. Technology and mobile health (mHealth) interventions can reduce barriers and improve access to diabetes education in rural communities. Screening, Brief Intervention, and Referral to Treatment (SBIRT) and financial incentives can be used with mHealth interventions to increase the uptake of diabetes education; however, studies have not examined their combined use for diabetes self-management in rural settings.

Objective: This two-phase Stage 1 feasibility study aimed to use a mixed methods design to examine the feasibility and acceptability of an mHealth diabetes education program combining SBIRT and financial incentives to engage rural individuals.

Methods: In Phase 1, we aimed to develop, adapt, and refine the intervention protocol. In Phase 2, a 3-month quasi-experimental study was conducted with individuals from 2 rural communities in South Texas. Study participants were individuals who attended free diabetes screening events in their community. Those with low or medium risk received health education material, whereas those with high risk or those with a previous diagnosis of diabetes participated in motivational interviewing and enrolled in the 6-week mHealth Diabetes Self-Management Education Program under either an unconditional or aversion incentive contract. The participants returned for a 3-month follow-up. Feasibility and acceptability of the intervention were determined by the rate of participant recruitment and retention, the fidelity of program delivery and compliance, and the participant's satisfaction with the intervention program.

Results: Of the 98 screened rural community members in South Texas, 72 individuals met the study eligibility and 62 individuals agreed to enroll in the study. The sample was predominately female and Hispanic, with an average age of 52.6 years. The feedback from study participants indicated high levels of satisfaction with the mHealth diabetes education program. In the poststudy survey, the participants reported high levels of confidence to continue lifestyle modifications, that is, weight loss, physical activity, and diet. The retention rate was 50% at the 3-month follow-up. Participation in the intervention was high at the beginning and dissipated in the later weeks regardless of the incentive contract type. Positive changes were observed in weight (mean -2.64, SD 6.01; $P<.05$) and glycemic control index (-.30; $P<.05$) in all participants from baseline to follow-up.

Conclusions: The finding showed strong feasibility and acceptability of study recruitment and enrollment. The participants' participation and retention were reasonable given the unforeseen events that impacted the study communities during the study

period. Combining mHealth with SBIRT has the potential to reach individuals with need to participate in diabetes education in rural communities.

(JMIR Mhealth Uhealth 2020;8(3):e16683) doi:10.2196/16683

KEYWORDS

Screening, Brief Intervention, and Referral to Treatment (SBIRT); Hispanic Americans; behavioral economics; rural population; diabetes; screening

Introduction

Background

Obesity and type 2 diabetes mellitus (T2DM) are becoming the most prevalent chronic illnesses in the United States and worldwide [1,2]. In the United States, over 30 million people, 9.4% of the US adult population live with diabetes and 23.8% of them are undiagnosed [3]. In Texas, 11.2% of the adult population has been diagnosed with diabetes [4]. Furthermore, the prevalence of diabetes is 17% higher in rural communities than urban communities [5], and it has been repeatedly recognized as the number 3 rural health priority [6,7]. Screening is essential to identify undiagnosed diabetes, allowing individuals to access resources to manage their illness. Screening, Brief Intervention, and Referral to Treatment (SBIRT) is an evidence-based practice that involves screening patients using a validated, standardized tool, referral for patients who need additional services to brief intervention with a health care professional, and referral to treatment [8]. Traditionally used to reduce alcohol and illicit drug use, SBIRT and its benefits can be realized beyond substance abuse, and these have been applied to undiagnosed hypertension [9] and childhood obesity [10-12]. For example, Byrne et al [12] demonstrated that an electronic health screening, a brief intervention, and referral to treatment within a primary care setting were effective to engage parents of overweight children in taking preventive actions. Financial incentives, based on the principles of behavioral economics, is another strategy that has demonstrated positive effects on the uptake of lifestyle modification and disease self-management [13]. Therefore, financial incentives may play a key role in engaging low-income and minority patients in changing habitual physical activity and dietary behaviors and diabetes management skills [14,15]. To the best of our knowledge, there are no published studies that examine the use of SBIRT and financial incentives to engage rural populations in diabetes self-management.

Research has indicated that behavioral healthy lifestyle interventions that focus on self-monitoring and goal setting are effective for glycemic control [16,17]. The landmark Look AHEAD (Action for Health in Diabetes) trial demonstrated that a lifestyle intervention could achieve clinically significant weight loss and glucose control in overweight and obese adults with T2DM [18]. Previous research has documented meaningful changes in glycemic control and other diabetes-related outcomes among rural patients who participated in diabetes self-management programs offered by their primary care providers in underserved rural communities [19,20]. However, novel approaches are still needed to increase the uptake of lifestyle diabetes prevention and management programs in underserved minority and rural populations [21-24]. A recent pilot study targeting underserved individuals demonstrated the positive effects (eg, weight loss) of using mobile versus paper-based logs for diabetes self-management, indicating that technology can be integrated into interventions to assist with self-monitoring [25]. Individuals can learn how to manage their disease through diabetes education provided by a trained leader. Diabetes education is also a cost-effective way to deliver training in self-management behaviors to individuals with T2DM [26]. Although diabetes education can have a significant impact on these individuals, nearly half of the adults with diabetes in the United States have not received formal diabetes education [27].

Resources for diabetes self-management are lacking in rural communities, where the availability of diabetes education programs is sparse. People from rural areas comprise 18% of the total US population, but rural areas account for 84% of the total area in the United States [28]. Although rural populations comprise a much smaller portion of the US population, as noted above, the prevalence rate for diabetes is 17% higher in rural areas compared with nonrural areas [29]. Furthermore, although there is a higher rate of diabetes in rural populations, there are fewer medical services available, and those services are more difficult to access. In particular, southern United States is the least likely to offer diabetes education [26]. Several studies have described the disparities that exist between rural and urban regions, with rural individuals living in medically underserved areas having older individuals, fewer employment opportunities, higher rates of underinsured and uninsured, lesser educational attainment, and being more likely to live in poverty [26,29,30]. Furthermore, even when rural individuals have access to diabetes self-management programs, participation may be limited because of the distance and transportation requirements to attend the group in person [31]. One solution to improve access to diabetes self-management programs in rural areas is to use technology to deliver education. Several studies have indicated the effectiveness of technology and mHealth to reduce barriers to diabetes self-management and improve health outcomes [25,31,32].

Study Objectives

This study combined mHealth, SBIRT, and the principles of behavioral economics to reach adults living in rural areas who were at risk for, or living with, diabetes. We tested the feasibility and acceptability of (1) recruiting rural residents to participate in diabetes screening, (2) motivating rural residents at risk for diabetes or with diabetes to make lifestyle modifications, and (3) engaging rural residents and encouraging them to complete an mHealth diabetes education program using different incentive plans. The long-term goal of the study is to develop an effective community-based diabetes education program for resource-poor

rural communities at increased risk for diabetes-related health disparities. Findings from this study and lessons learned will guide future health intervention research with rural populations.

Methods

Study Design

This was a two-phased Stage 1 feasibility study according to the National Institutes of Health's staged model for the development of psychosocial and behavioral interventions [33]. Using a mixed methods design, we collected both qualitative and quantitative data to evaluate the aims of the study. The study team used Phase 1 of the study (Stage IA) to identify community partners, develop recruitment plans, and adapt an integrated diabetes screening and brief intervention and motivational interviewing to a diabetes education program, with input from the community stakeholders. During Phase 2 of the study (Stage IB), we assessed the feasibility and acceptability of the diabetes education program using a quasi-experimental pretest and posttest design. The aims of Phase 2 were to gather data on the rates of study enrollment, recruitment and retention, delivery of mHealth education content, levels of engagement with the scheduled activities under either unconditional or aversion incentive contract, acceptability of and satisfaction with the diabetes screening and education program, and the sensitivity of outcomes (weight and glycemic control index) to the mHealth intervention.

Phase 1 of the Study: Program Development and Adaptation for Delivery

During Phase 1 of the study, the study team conducted numerous site visits and successfully identified 2 community hospitals and a faith-based community nursing program as partners for conducting free diabetes screening and recruiting study participants. There was no ongoing program that offered free diabetes screening or diabetes prevention or management programs in these communities. On the basis of the information gathered from the 2 study communities, the study team developed the pilot version of the diabetes screening and education program by adapting SBIRT, financial incentives, and the mHealth diabetes education program from a previous project. We offered the pilot version of the program to 6 residents recruited from a local community. The participants completed the diabetes screening and received brief diabetes education based on their screening results. The study staff also explained the content of an mHealth education program to the participants. At the end, the study team conducted a focus group discussion with the participants to gather their feedback on the recruitment, procedure of diabetes screening and brief education, and content and delivery of the mHealth education program. The study team refined and finalized the intervention and data collection protocol based on the finding from the focus group discussion.

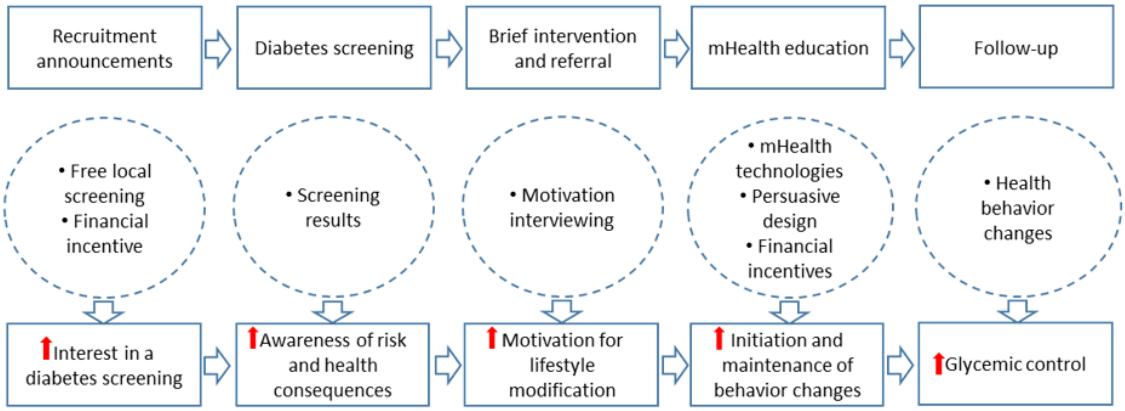
Participant Recruitment in Phase 2 of the Study

Study participants lived in 2 rural communities in South Texas (Community A and Community B). Diabetes screening events were promoted by distributing recruitment flyers in community hospitals, in supermarkets, in churches, and at community events, and announcements were made on local radio stations and newspapers. To qualify for the study, participants needed to meet the following criteria: (1) live in a study community, (2) be 35 years or older, (3) own a smartphone (with a data plan) or a mobile phone (with an SMS text message plan) and have internet access via a computer or tablet, (4) complete a diabetes screening, and (5) be at high risk for diabetes, scoring >5 on the American Diabetes Association (ADA) diabetes risk test or A_{1c} level ≥5.8, or have health care provider–diagnosed diabetes. English language proficiency was not an inclusion criterion. Residents who did not meet the eligibility criteria only received diabetes screening and brief education and were not eligible to receive the study incentive. Residents who met the eligibility criteria but had low or moderate risk for diabetes or were not interested in the study received brief health counseling and health promotion materials. All eligible participants who completed diabetes screening received a US $25 grocery store gift card, and participants who completed the mHealth diabetes education program received up to US $60 in grocery store gift cards.

Description of the Intervention: Diabetes Screening and Mobile Health Diabetes Education Program

The intervention was based on the principles of SBIRT and behavioral economics for program participants who were ready to commit to lifestyle modification with minimal social and technical support. The principles of persuasive design were used to guide the development of the mHealth diabetes education program that stresses ease of access and motivation [34]. To shed light on the intervention approach, Figure 1 shows the conceptual framework that depicts the underlying processes to engage the study participants. We conducted the Diabetes Screening and mHealth Education Program following SBIRT [8] in 3 consecutive steps, which are described in the following sections.

Figure 1. Conceptual Framework of the Diabetes Screening and mHealth Education Program.



Step 1: Free Diabetes Screening

The screening was conducted by trained research assistants either in the 38-foot Mobile Health Laboratory or a large indoor space within the community. The diabetes screening included weight assessment and calculation of BMI, resting blood pressure, blood assay by finger stick for hemoglobin A_{1c}, high-density lipoprotein cholesterol, total cholesterol, and glucose. Each screened participant was assessed for the risk of type 2 diabetes based on the multiple risk factor diabetes screening and education guidelines of the ADA and the US Preventive Services Task Force. Low risk was defined as scoring <5 on the ADA diabetes risk test [35], moderate risk was defined as scoring 5 on the ADA diabetes risk test, and high risk was defined as scoring >5 on the ADA diabetes risk test or an A_{1c} level ≥5.8. Participants with diabetes self-reported diagnosis of diabetes by a physician. The results of the screening were recorded on a screening report card. All participants who agreed to participate in the study were invited to complete a health needs survey and received a study information sheet. The institutional review board reviewed and approved all study protocols.

Step 2: Brief Intervention and Referral to Treatment

The screened residents were divided into 4 groups based on the results of the diabetes screening and willingness to participate in the study (Table 1). The study participants in groups 1 and 2 received brief education by trained research staff members (a certified community health worker or registered dietitian) who provided an explanation of screening results and brief counseling on healthy lifestyle behaviors. Each participant was provided health education brochures on various topics related to general healthy lifestyle behaviors and diabetes prevention.

Table 1. Grouping of study participants and participant treatment.

Screening result	Brief intervention	Referral to treatment	Incentives
Group 1: Not meeting eligibility or not interested in the study	Brief education by a community health worker (5-10 min)	No referral	• US $25 grocery card for participating in screening
Group 2: low-to-medium risk for diabetes	Brief education by a community health worker (5-10 min)	No referral	• US $25 grocery card for participation in screening
Group 3: high risk for diabetes	Motivational interviewing by trained research staff (15-30 min)	Mobile health Diabetes Education Program	• US $25 grocery card for participation in screening • Unconditional incentives (US $60) for completing a diabetes education program
Group 4: previously diagnosed diabetes	Motivational interviewing by trained research staff (15-30 min)	Mobile Health Diabetes Education Program	• US $25 grocery card for participation in screening • Aversion incentives (up to US $60) for completing a diabetes education program

The participants in groups 3 and 4 received brief motivational interviewing based on the Feedback, Responsibility, Advice, Menu Options, Empathy and Self-Efficacy approach that includes *Feedback* regarding demographic and biological risk for diabetes, emphasis on personal *Responsibility* and choice, *Advice* to change (when appropriate), a *Menu* of change options, an *Empathetic* listening approach, and an emphasis on *Self-efficacy* and optimism around change [36]. Research staff (faculty researchers, registered nursing students, and a registered dietitian) completed a 5-hour motivational interviewing training session provided by a clinical psychologist with expertise in SBIRT and motivational interviewing (MI). Trained research staff used a personalized screen report to guide the delivery of the brief MI intervention content. Section 1 of the screen report provided a health risk profile and associated health consequences based on the screening results [37]. Section 2 demonstrated the

risk reduction and improvement of glycemic control based on the research evidence of lifestyle modifications (weight loss, diet, and physical activity). Section 3 included a menu of options for behavior changes and evidence-based strategies [38]. With facilitation from the research staff, participants identified behaviors deemed important to them for modification and selected goals and strategies of lifestyle change, which they were willing to attempt to adopt for the future. At the end of the brief motivation interviewing session, participants were asked to indicate their level of readiness on a scale of 0 to 10 for making lifestyle changes at the present time.

The participants with a readiness score of 5 or higher were asked if they were willing to complete a 6-week mHealth diabetes education program. Once the participants agreed to enroll in the program, a research staff member explained the details of the mHealth education program and the weekly schedules of activities. Thereafter, the study participants signed a study contract to indicate their willingness to participate in the study and complete the study activities. The content of the study contract included the following: (1) commitment to follow the schedule of program activities and (2) agreement to receive a monetary incentive for following the schedule of program activities. All participants in community A received a US $60 grocery store card after they signed the contract (unconditional incentive group). All participants in community B were promised a grocery store card worth US $60 if they completed all scheduled activities, reductions would be made for each scheduled activity that was not completed (aversion incentive group). Finally, the study participants received an information packet, which included the mHealth diabetes education program schedule, the study incentive tracking form (for community B), and study staff contact information.

Step 3: Delivery of the Mobile Health Diabetes Education Program

The program was delivered over a 6-week period with a different topic area for each week (see Table 2). Weekly interactive lessons were created for the participants to develop lifestyle modification knowledge and skills based on the National Diabetes Education Program and best practices for diabetes self-care [39,40]. We used Articulate Storyline (Articulate Global Inc) to develop interactive education lessons, with avatars, problem-solving quizzes, skill-building games, and embedded video clips to motivate and engage participants. At the end of each lesson, participants watched a multi-part video drama of a family dealing with diabetes. Each week, the participants were asked to complete one physical activity challenge and one diet challenge out of the 2 to 3 challenges offered, which were related to the skills presented in the lessons. Resources relevant to the topic of the week were provided to participants to further their understanding of the topic. These included YouTube videos and websites with information on physical activity, nutrition, and stress reduction. The participants accessed the content via their smartphone or internet-connected computers or tablets. Automated SMS text messages were used to send program reminders and health tips relevant to the topic of the week. The SMS text messages were delivered to participants using a reconfigurable SMS text messaging system, MessageSpace that used the “grouping” logic to define various cohorts and schedule the delivery of SMS text messages for a comparative analysis or send group-specific broadcast messaging, “polling” for collecting responses from participants, and advanced message history tracking and scheduling to determine dose/exposure.

Table 2. Mobile Health Diabetes Education Program curriculum.

Diabetes education lesson (to be completed by Wednesday)	Health challenge (to be completed by Sunday)	SMS text messages/polling (review or respond upon receiving)	Resources (review by Sunday)
Week 1: Understanding diabetes and obesity; eating healthy to manage or prevent obesity and diabetes	1 physical activity challenge and 1 diet challenge	2 program reminders and 2 daily texts	Videos on physical activity; 1 video on diabetes risk and consequences; and 1 website with information on diabetes risk and health consequences
Week 2: Understand what foods go in a healthy lifestyle; understand portion control and moderation	1 physical activity challenge and 1 diet challenge	2 program reminders and 2 daily texts	Videos on physical activity; 1 video on healthy eating; and 1 website with information on healthy eating strategies
Week 3: Learning important nutrition terms; learn how to read a nutrition label; and learn what is healthy vs unhealthy	1 physical activity challenge and 1 diet challenge	2 program reminders and 2 daily texts	Videos on physical activity and 1 website with information on nutrition facts
Week 4: Learn what counts as physical activity; learn how physical activity helps diabetes; choose an activity that is fun for you; and class stretching activity	1 physical activity challenge and 1 diet challenge	2 program reminders and 2 daily texts	Videos on physical activity; 1 website with information on physical activity
Week 5: Benefits of stress reduction; recognize symptoms of depression; how to reduce stress; importance of socializing; how to meet new people; and benefits of enough sleep	1 physical activity challenge and 1 diet challenge	2 program reminders and 2 daily texts	Videos on physical activity; 1 video on depression; 1 website with information on stress management; 1 breath control exercise video
Week 6: Diabetes myths; create a healthy plate for managing diabetes; and exercise and food for managing diabetes	1 physical activity challenge and 1 diet challenge	2 program reminders and 2 daily texts	Videos on physical activity; 1 video on diabetes management; 1 website with information on diabetes self-management

At the beginning of each week (Monday), the participants received an SMS text message to notify them about the educational activities for the upcoming week and a link to a REDCap portal that provided access to all educational content. An email with the same information was also sent to the participant's email address. The participants were asked to submit a report on the REDCap portal to the research team to acknowledge the receipt of class activities and report any technical issues. At the end of each week (Sunday), participants received an SMS text message to indicate whether they had completed the activities and encourage them to complete the activities if they had not done so. If there was no indication of participation for more than one week, a research staff member called the participants and asked whether they needed help. The participants were also encouraged to call the study team if they needed technical assistance.

After completing the 6-week program, participants received weekly SMS text messages with health tips for an additional 6 weeks. The study participants returned to participate in a follow-up screening event (ie, the posttest) at the end of the 3-month study period. They were debriefed on the study and received additional information on lifestyle modification strategies and local health resource maps to support their continued effort of lifestyle change. The participants in the aversion group also received their grocery card incentive.

Measurement and Evaluation

Evaluation of the Feasibility of Study Implementation

The information used to assess the feasibility of study implementation included (1) characteristics of the residents interested in the study, (2) feasibility of recruitment strategies and participation and retention of study participants, (3) participant engagement and compliance with the intervention, (4) fidelity of intervention delivery, (5) standardization of the intervention protocol, (6) refinement of outcome measurement and process evaluation protocols, as well as staff training program, and (7) completion rate of biometric measures.

Evaluation of Acceptability of the Intervention

After completing the follow-up diabetes screening, the participants were invited to complete an anonymous survey to obtain their perception of the program benefits, satisfaction with program activities, and evaluate their confidence to continue efforts to maintain a healthy lifestyle, and a 15-to-20-min debriefing interview was conducted with a faculty researcher.

Biometric Data

All participants completed 2 diabetes screenings that included biometric information at baseline and 3-month follow-up. Data were collected following a standardized protocol by trained research assistants in the University of Texas at San Antonio Mobile Health Laboratory, in a 38-foot customized recreation vehicle, or within the space provided by the study partners in a community setting. The protocol was also piloted using both the Mobile Health Laboratory and indoor space to identify the most efficient use of space and staff. Height and weight were measured twice, and participants wore light clothing. BMI (in kg/m^2) was calculated by taking the average of two measures, which required discrepancies of less than 0.5 kg and 0.5 cm for weight and height measurements, respectively. Hemoglobin A_{1c} (A_{1c}) was assessed by using a finger-stick testing machine (A1CNow+ System, PTS Diagnostics). In a randomized controlled trial (RCT), BMI for participants at risk for diabetes and A_{1c} for participants with diabetes would be the primary outcome measures to assess the efficacy of the intervention program. Resting systolic and diastolic blood pressure was measured using an electronic blood pressure monitor after a 5-min rest. Three measures were taken; the two closest measures were averaged. We did not include the blood pressure data in the analysis as the testing environment varied from baseline to follow-up screening.

Data Analysis

Descriptive statistics were calculated for feasibility as well as demographic and biometric data by community location (incentive groups) and for the total sample. Furthermore, *t* tests (two-tailed) for continuous variables and chi-square tests for categorical variables were used to determine the demographic and biometric differences between the incentive groups over time (baseline to follow-up). Analysis of variance was performed to test the changes in the biometric data from baseline to follow-up, controlling for significant demographic covariates ($P>.25$). Analyses were performed in SPSS version 23 (IBM Corp, 2017) and Stata version 14 (Stata Corp, 2017). Content analysis was used by a faculty researcher to analyze the participant responses in the follow-up interviews [41].

Results

Phase 2: Study Recruitment and Program Enrollment

The diabetes screening protocol was designed to resemble diabetes screenings offered in community or worksite health fairs. The quality of the screening was enhanced by offering a brief intervention and referral to treatment based on the screening outcomes. The characteristics of the residents attending the screening events and the study sample in the 2 study communities are shown in Table 3. The screening goal was 100 residents who would attend the free screening events. We conducted 6 screening events in the 2 communities (3 in community A and 3 in community B), with 98 residents attending the events. A total of 72 of these residents met the preliminary eligibility and completed some or all of the diabetes screening. A total of 62 of the screened residents (28 in community A, the unconditional incentive group, and 34 in community B, the aversion incentive group) enrolled in the mHealth diabetes education program and became study participants (yield rate of 67.4%). There was no difference in demographic and biometric variables between screened and study participants. Over 70% (22/79) of the participants were female, primarily Hispanic, had a family history of diabetes, and accessed the internet with a smartphone. A total of 69% of the participants were 45 years old or older. Less than 6% of the participants reported speaking Spanish at home or using Spanish for reading. Over 25% of the participants did not have a primary care provider. The average A_{1c} was 6.9 (SD 3.4) and 6.5 (SD 2.0) at baseline for the screened participants and study

participants, respectively. The demographic and health characteristics of the study participants in community A and community B were similar, with the exception that the rate of diagnosed diabetes (P<.01) and BMI (P<.05) were higher in the latter. Figure 2 shows the flow of the study participants.

Table 3. Characteristics of study sample and screened residents.

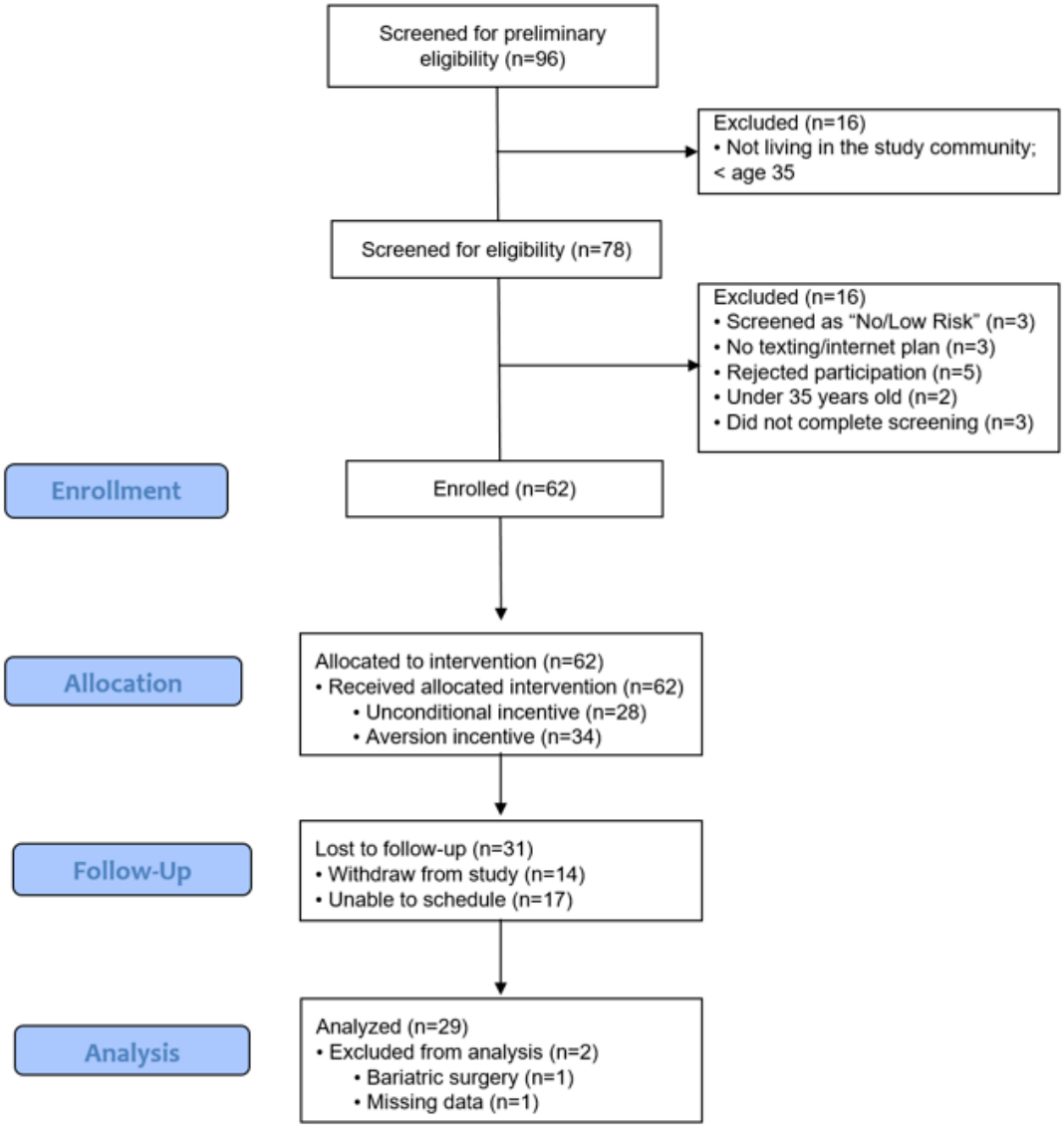
Variable[a]	Unconditional incentive (n=28)	Aversion incentive (n=34)	Total study sample (n=62)	Attendants of all events (n=78)
Female gender, n (%)	22 (79)	23 (68)	45 (73)	54 (69)
Hispanic, n (%)	21 (75)	20 (74)	46 (74)	55 (71)
Family history of diabetes, n (%)	20 (71)	29 (85)	49 (79)	58 (75)
Gestational diabetes, n (%)	3 (14)	7 (30)	10 (22)	14 (18)
Diabetes diagnosis[b], n (%)	3 (11)	14 (41)	17 (27)	24 (31)
Participated with others, n (%)	12 (43)	15 (44)	27 (44)	27 (35)
Owning a cell phone, n (%)	27 (96)	33 (97)	60 (97)	73 (94)
Having an email, n (%)	22 (79)	24 (71)	46 (74)	53 (68)
Language spoken at home, n (%)				
English	25 (89)	29 (85)	54 (87)	65 (83)
Spanish	1 (4)	3 (9)	4 (7)	4 (5)
Both	1 (4)	1 (3)	2 (3)	5 (6)
Other language	0 (0)	0 (0)	0 (0)	0 (0)
Language used for reading, n (%)				
English	25 (89)	29 (85)	54 (87)	65 (83)
Spanish	1 (4)	3 (9)	4 (7)	4 (5)
Both	1 (4)	1 (3)	2 (3)	2 (3)
Other language	0 (0)	0 (0)	0 (0)	0 (0)
Having a primary care provider	22 (79)	24 (71)	46 (74)	55 (71)
Age (years), mean (SD)	53.9 (11)	51.5 (11)	52.6 (11)	52.7 (12)
Weight (lbs), mean (SD)	194.8 (33)	216.4 (58)	207.4 (50)	206.7 (49)
BMI (kg/m^2)[c], mean (SD)	30.4 (10)	36.6 (8)	34.1 (9)	34.5 (8)
Total cholesterol, mean (SD)	187.9 (64)	165.5 (34)	173.3 (47)	173.3 (47)
High-density lipoprotein, mean (SD)	52.7 (16)	46.8 (16)	48.8 (16)	48.8 (16)
Hemoglobin A_{1c}, mean (SD)	6.2 (2)	7.4 (4)	6.9 (3)	6.5 (2)

[a]Chi-square test for categorical variables and independent *t* test for continuous variables for comparison of the treatment groups.

[b]P<.01.

[c]P<.05.

Figure 2. Study Participant Flow.



Phase 2: Participant Engagement and Retention

The mHealth Diabetes Education Program was delivered to the participants in 4 cohorts starting in late July 2017. Hurricane Harvey, a Category-4 hurricane (August 17,2017-September 2, 2017) landed on the Texas coastline and caused historic flooding and interruptions in mobile phone and internet services in South Texas, including the study communities. During this period, 14 participants dropped out of the study. Although we were not certain to what extent this inopportune event impacted the participants' participation in the mHealth Diabetes Education Program, the expectation that the study participants would adhere to the scheduled intervention activities seemed highly inappropriate. The impact was severe in the later cohorts. This prevented us from being able to evaluate the effects of the financial incentives on the participant's engagement in the program. To gauge the levels of participation, we tracked the participants' responses to the scheduled activities each week as indicators of intervention compliance. Table 4 shows the number of weeks (mean 2.5, SD 2.31 weeks, for the total sample) the participants responded to the intervention activities, which is measured by a record of a reported participation in at least one activity from the REDCap portal, a reply to an SMS text message, or telephone contact with the participant.

Table 4. Number of weeks the participants responded to the intervention activities by treatment group.

Number of weeks responded	Unconditional incentive (n=28)	Aversion incentive (n=34)	Total sample (N=62)
0, n (%)	5 (18)	10 (29)	15 (24)
≥1, n (%)	23 (82)	24 (70)	47 (76)
≥2, n (%)	19 (68)	21 (61)	40 (65)
≥3, n (%)	15 (54)	19 (55)	34 (55)
≥4, n (%)	11 (39)	17 (50)	28 (45)
≥5, n (%)	9 (32)	14 (41)	23 (37)
≥6, n (%)	6 (21)	7 (21)	13 (21)
Average of 6, mean (SD)	2.5 (2.2)	2.4 (2.3)	2.5 (2.1)

An additional analysis showed that the persons who participated in the study with relatives or friends (mean 3.2, SD 2.35) participated more often than those who participated in the study alone (mean 2.0, SD 1.92; t_{60}=2.17; P<.03). Table 5 shows the response rate to the intervention activities, indicated by how many participants participated in the intervention by week. The average weekly response rate of 4 cohorts was 41.4% for the total sample. There were no discernible differences by cohort or between the 2 incentive groups.

The 12-week follow-up screening was planned as a celebration of program completion and evaluation of progress in lifestyle modification. In total, 31 of 62 study participants (response rate: 50%) attending the follow-up screening. During the first week of the follow-up screening, a mass shooting killed 26 adults and children and wounded 20 others in a church less than 30 miles away from the 2 study communities. We decided to terminate the follow-up screening, which seemed inappropriate given the social and psychological impact of the shooting on the community. We attributed the low retention rate partly to this shooting event as well as the occurrence of a hurricane early in the intervention phase.

Table 5. Response rate to the intervention activities by week by the treatment groups

Study week	Unconditional incentive (n=28), n (%)	Aversion incentive (n=34), n (%)	Total sample (N=62), n (%)
1	19 (67)	24 (71)	43 (69)
2	12 (43)	13 (38)	25 (40)
3	11 (39)	14 (41)	25 (40)
4	13 (46)	11 (32)	24 (38)
5	9 (32)	11 (32)	20 (32)
6	6 (21)	11 (32)	17 (27)
Average of 6 weeks	12 (42)	14 (41)	26 (41)

Feasibility and Acceptance of Diabetes Screening Protocol in Phase 2 of the Study

The screening protocol was efficient to screen the number of participants, with a team of 5 trained research assistants for collecting the biometric data, 2 community health workers for delivering brief education to low-to-moderate risk residents, and 2 or 3 trained research staff for motivational interviewing with high-risk and diabetic participants. The use of the Mobile Health Laboratory offered a standardized environment to collect quality data for evaluation of the program's impact on glycemic control and body weight. To maintain high reliability and reduce testing interferences from the surrounding environment and weather, all biometric data, especially finger-stick tests and blood pressure, were collected from the Mobile Health Laboratory.

The completion rate of weight and A_{1c} measurement was 93.5% at baseline and 93.6% at 12-week follow-up. The main reasons for missing the weight and A_{1c} were the failure to collect data from the diabetes screening report and to run the blood assay using finger-stick tests. Some blood pressure measurements were not collected. A total of 38% (24/62) of the participants at baseline did not fast overnight. The completion rate of blood pressure and lipid measure was less than 70%. As a result, blood pressure and fasting glucose were not included in the analysis. There was no adverse event during the screening events. A total of 5 of 17 (29.4%) participants with A_{1c} ≥6.5 had not been diagnosed previously as having diabetes and were referred to treatment at the 2 community hospitals.

Fidelity of Implementation of Diabetes Screening and Mobile Health Education Program in Phase 2 of the Study

All program activities were implemented as scheduled by the research team. The REDCap portal worked effectively in presenting the program activities to the participants in a standardized format. There was no report of difficulty in navigating the content on the REDCap portal. However, the study participants in their follow-up interviews reported that

weak smartphone signals and unstable internet download speed caused difficulty in completing the interactive health lessons and interrupted their viewing of the videos, and others complained of not being able to see the content in the REDCap portal because of formatting issues on their phone. Records from MessageSpace showed that out of the 85 SMS text messages delivered, 49% of the participants received all of the SMS text messages, 43% of the participants received 75% of the texts, and 8% of the participants received ≤50% of the texts. Anecdotally, the participants reported difficulty in understanding the instructions for completing the activities; therefore, they often did not submit the report on the REDCap portal even if they had reviewed the content.

Acceptability of Diabetes Screening and Mobile Health Education Program in Phase 2 of the Study

Table 6 shows the results of the poststudy survey that participants completed at the follow-up screening. The participants reported that the program helped them to be more active than before, eat healthy, and lose weight, with 73.6% to 94.1% of the participants indicating responses from "agree" to "strongly agree." The participants reported high levels of satisfaction with the program content (interactive lessons, SMS text messages, and videos). They indicated confidence to continue lifestyle modification with >85% indicating responses from "agree" to "strongly agree." Finally, the participants were asked to indicate if exercising regularly, eating healthy, or losing weight is now an essential, high, moderate, low, or no priority activity compared with the beginning of the program. The majority of the participants indicated that exercising regularly (73%), eating healthy (74%), and losing weight (77%) had become essential or high priority (data not shown). Overall, the participants were least satisfied with the weight loss component of the program.

Table 6. Poststudy survey of the diabetes education program (n=25).

Questions	Strongly agree, n (%)	Agree, n (%)	Disagree/strongly disagree, n (%)
Please answer these questions as honestly as possible.			
Did the eDiabetes Education Program help you to be more physically active?	8 (32)	14 (55)	3 (12)
Are you still being active with the information from the eDiabetes Education Program?	10 (38)	13 (53)	2 (9)
Did the eDiabetes Education Program help you to eat healthy?	10 (38)	14 (56)	1 (6)
Are you still eating healthy with the information from the eDiabetes Education Program?	8 (32)	15 (59)	2 (9)
Did the eDiabetes Education Program help you to lose weight?	6 (24)	12 (50)	7 (26)
Are you still trying to lose weight with the information from the eDiabetes Education Program?	7 (27)	14 (56)	4 (18)
I liked the weekly diabetes education lessons.	8 (32)	17 (68)	0 (0)
I learned how to change my lifestyle with information from the health education lessons.	9 (35)	15 (59)	1 (6)
I liked the weekly text messages with health tips.	12 (49)	12 (49)	1 (3)
I liked the weekly health challenges.	9 (34)	15 (60)	1 (6)
I liked the YouTube videos on physical activity and diet.	8 (30)	14 (58)	3 (12)
I liked the YouTube videos with information on obesity and diabetes.	8 (30)	13 (55)	4 (15)
Compared with when the eDiabetes Education Program started in the summer...			
I am confident that I can continue to exercise regularly	11 (43)	13 (51)	1 (6)
I am confident that I can continue eating healthily	10 (40)	14 (54)	1 (6)
I am confident that I can continue to lose weight	11 (43)	11 (43)	3 (14)

The themes generated from the follow-up interviews with 13 participants are shown in Table 7. The participants perceived that the program was motivational and had positive effects on their lifestyle modification effort. Their experience was negatively influenced by technical issues with slow internet speed, weak smartphone signals, and lack of Web-based support. They encountered various nontechnological barriers to complete the program activities, such as family commitment and work conflicts. The impact of the hurricane also prevented some from continuing the program. Respondents stated that the program could be improved by offering some human interactions with the participants, improving the interface on REDCap portal to reduce participant burden, and providing training on "how to do things."

Table 7. Themes of participant interviews and exemplar statements (n=13).

Themes	Number of responses	Exemplars
Positive experiences with the program	38	• "I have eaten more fruit than I ever have, watch my soda intake and eat steamed vegetables. My wife cooks differently now" • "Changing my diet and walking. Tried to eat more healthy meals. Helpful to do with the family" • "I did the challenges. When I first did the challenges, I picked 20 minutes. Now it is nothing. I used to have no energy and now I do"
Motivators	19	• "I think if it were not for the texts message, I would have let go long ago" • "I knew I wasn't eating well, I wanted to change my diet and my husband's" • "Starting this motivated me. When I came here [for the initial screening], my weight was so high!"
Technological issues	20	• "The video kept freezing. It was really frustrating" • "When I got the texts, I could not see all of it. It was frustrating…I was trying so hard" • "Had it on the phone. More adapted to phone. Had a lot of trouble with computer" • "The videos were the least useful… basic and banal…hard to watch"
Circumstances undermining the ability to engage as anticipated	10	• "The hurricane damaged my home…hard to catch up. I have custody of my grandchildren. I am always working and tired" • "At the beginning I tried to do all the challenges, towards the end I wasn't. Summer vacation was over, when I went back to work, I could not do anymore. Beginning of the school year is a stressful time" • "This got put on the back burner, I was working fulltime, going to school, two children, and we were moving"
Suggestions to improve the program	11	• "Have activities in town, especially for senior citizens. Face-to-face exercising with a group" • "Have a more user-friendly way of showing participants what challenges they have and have not completed" • "Have people keep a log, food journaling" • "Improve how we submit our progress online

Sensitivity of the Study Outcome Measures to the Diabetes Screening and Mobile Health Education Program

Table 8 displays the unadjusted changes in the study outcomes from baseline to follow-up between the incentive groups and the total sample. Overall, all of the outcome measures showed positive responsiveness to the intervention with a significant reduction in weight (P<.05) in the total study sample from baseline to follow-up over a 12-week period (see Table 8).

Table 8. Unadjusted changes of the outcome measures from baseline to follow-up by treatment group and for total study sample.

Variable[a]	Unconditional incentive		Aversion incentive		Total sample	
	n	Mean (SD)	n	Mean (SD)	N	Mean (SD)
Weight (lbs)	13	–3.88 (3.44)	16	–1.63 (7.46)	29	–2.64[a] (6.01)
BMI (kg/m^2)	13	–0.24 (1.53)	16	–0.03 (1.62)	29	–0.12 (1.55)
Total cholesterol	6	–42.33 (116.97)	13	–2.15 (26.44)	19	–14.84 (68.08)
High-density lipoprotein	6	4.71 (18.20)	14	3.45 (7.82)	20	3.89 (11.98)
Hemoglobin A_{1c}[b]	13	0.18 (0.65)	16	–0.68 (1.29)	29	–0.30 (1.12)

[a]The analysis of variance for comparing the changes between the treatment groups; paired t test for comparing the changes in the total sample from the baseline to the posttest.

[b]P<.05.

Interferences in participants' responses to the scheduled activities due to hurricane-related flooding and small sample size due to low retention made the comparison by the incentive group not meaningful. Therefore, the effects of the 2 incentive plans on the outcome measures including a significant reduction in A_{1c} in community B compared with that in community A should be interpreted with caution.

Discussion

Principal Findings

Findings from this feasibility study lend support to the conceptual framework of the intervention and the feasibility and acceptability of combining SBIRT, the components of behavioral economics, and an mHealth diabetes education in underserved rural communities. The protocol for recruitment and program enrollment was successful in generating the targeted study sample, but the retention rate can be improved through the lessons learned. Using both qualitative and quantitative data, the study demonstrated high levels of satisfaction and acceptability of free diabetes screening and brief counseling among the rural residents. In light of numerous challenges encountered in the delivery of the mHealth content, participants' responses to intervention activities were mostly favorable. The study team gained important insights to address the barriers and improve the mHealth intervention for this underserved rural population.

The findings of the study supported the premise and conceptual framework of the intervention. The participants reported high levels of satisfaction with the program content as well as increases in their abilities and priorities to improve physical activity, diet, and weight loss as a result of the intervention. Satisfaction with the program and changes in behavioral intentions are important mediators and moderators of program effectiveness and are associated with positive behavior change outcomes [42,43].

It was important to this study to determine the feasibility of recruiting and engaging rural residents of 2 communities with largely Hispanic populations into an mHealth diabetes prevention program. The feasibility of recruitment was demonstrated by screening 96 residents within the 6 scheduled recruiting events. Of those screened, 62 enrolled in the study for an enrollment rate of 65%, which is comparable with other studies engaging rural Hispanic individuals [28-30]. However, this study had a low retention rate (50%) compared with a similar mHealth diabetes intervention conducted with a disparate population [19]. The high rate of attrition in this study may be attributed to the ramifications of Hurricane Harvey, which occurred in August 2017, and the Sutherland Springs shooting, which occurred within 30 miles of the data collection site, during the time this pilot study was conducted [31].

Although enrollment rates were sufficient for this two-phase study, future studies will require higher enrollment and retention to achieve sufficient power. Research has shown that word-of-mouth recruitment is an effective method for recruiting rural Hispanic populations, particularly Spanish-speaking first-generation immigrants, to participate in research [44]. Others have also shown that utilizing members of the community increased the rapport between recruiters and participants and should be considered as methods to increase enrollment and retention in future studies [45,46]. Furthermore, retention may be increased by using community health workers (CHWs) or promotoras. A CHW or promotora is a frontline public health worker who is a trusted member and/or has an unusually close understanding of the community served. This trusting relationship enables the CHW to serve as a link between health services and the community to facilitate access to services and improve the quality and cultural competence of service delivery [35]. A CHW also builds individual and community capacity by increasing health knowledge and self-sufficiency through a range of activities such as outreach, community education, informal counseling, social support, and advocacy [35]. Previous studies have demonstrated that retention in diabetes self-management programs is higher when promotoras are used compared with lay leaders [8] and that promotoras are trusted and provide social support in a different way than does a traditional provider [9]. As our study focused on electronic communication to engage participants, a hybrid "blended" model that utilizes CHWs to screen and deliver brief interventions, as well as provide follow-up throughout the study, has the potential to increase engagement and retention for the duration of an mHealth study [36].

The low response rate to the scheduled intervention activities cannot be fully attributed to differences in incentives in consideration of the interruptions from the hurricane-related flooding in the community. Furthermore, programs that emphasize knowledge acquisition without including behavior change have proven to be less effective in obtaining positive outcomes. There is agreement that education in itself does not lead to behavior change, as there are many factors both internal and external to an individual that lead to changes in health behaviors [10]. For example, goal setting has led to the attainment of behavior change goals for individuals with diabetes [11], and interventions focused on self-management that include both education and behavior change techniques have proved to lead to improvements in health outcomes, such as weight loss in individuals with diabetes [19]. Social support has been identified as a critical factor in mHealth interventions [19]. In this study, participants who enrolled in the program with a friend or family member had a higher response rate compared with those who participated alone. As such, to successfully change behaviors and obtain health outcomes, it will be necessary in future studies to focus on multiple factors that lead to behavior adoption and maintenance.

Limitations

There are several limitations that should be considered when interpreting the findings of this project. First, we collected data from a convenience sample and the findings may not generalize to adult residents of rural areas in the United States. Second, we did not collect process data to examine the usability of the platform delivering the intervention content, which could provide information on the participant's acceptability of the intervention [47] Third, study participants' health insurance status was not collected, this information could have strengthened the study findings and better informed the targeted

intervention based on insurance status in the future. Fourth, due to the exploratory nature of this project and the focus of the study on feasibility and acceptability, the statistical findings need to be interpreted with caution. Additionally, we had high attrition between the 2 time points in part due to 2 local disasters during the study period. Interpretations of study findings need to consider the impact of these 2 local disasters.

Conclusions

This pilot study demonstrated strong feasibility in recruiting rural patients with diabetes and delivering a technology-based lifestyle intervention, despite numerous challenges. The acceptability of the program was high. The participants perceived the program was motivational and had positive effects on their lifestyle modification effort. The feasibility and acceptability information collected from this mixed methods study generated meaningful insights for planning a large scale RCT of this intervention and suggested adding a human touch, for example, through CHWs. Future studies should examine the mechanisms of the mHealth intervention that influence the outcomes in diabetes control in this underserved rural diabetes population.

Acknowledgments

The authors would like to thank the participants for their time and community partners for their support. They also wanted to thank Dr David Akopian and his team at the University of Texas at San Antonio for their technical assistance in the use of MessageSpace. The funding support of the project was provided by the San Antonio Life Sciences Institute and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant UL1 TR001120. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004 May;27(5):1047-1053. [doi: 10.2337/diacare.27.5.1047] [Medline: 15111519]
2. Poirier P, Giles TD, Bray GA, Hong Y, Stern JS, Pi-Sunyer FX, et al. Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American Heart Association Scientific Statement on Obesity and Heart Disease from the Obesity Committee of the Council on Nutrition, Physical Activity, and Metabolism. Circulation 2006;113(6):898-918. [doi: 10.1161/circulationaha.106.171016]
3. Centers for Disease Control and Prevention. Atlanta, GA: Centers for Disease Control and Prevention; 2017. National Diabetes Statistics Report URL: https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf [accessed 2020-03-10]
4. Centers for Disease Control and Prevention. Atlanta, GA: Centers for Disease Control and Prevention; 2015. United States Diabetes Surveillance System URL: https://gis.cdc.gov/grasp/diabetes/DiabetesAtlas.html [accessed 2020-03-10]
5. Keppel K, Pearcy J, Klein R. Measuring progress in Healthy People 2010. Healthy People 2010 Stat Notes 2004 Sep(25):1-16. [Medline: 15446274]
6. Gamm L, Hutchison L. Rural healthy people 2010--evolving interactive practice. Am J Public Health 2004 Oct;94(10):1711-1712. [doi: 10.2105/ajph.94.10.1711] [Medline: 15451738]
7. Bolin JN, Bellamy GR, Ferdinand AO, Vuong AM, Kash BA, Schulze A, et al. Rural healthy people 2020: new decade, same challenges. J Rural Health 2015;31(3):326-333. [doi: 10.1111/jrh.12116] [Medline: 25953431]
8. Babor TF, McRee BG, Kassebaum PA, Grimaldi PL, Ahmed K, Bray J. Screening, brief intervention, and referral to treatment (SBIRT). Substance Abuse 2007;28(3):7-30. [doi: 10.1300/j465v28n03_03]
9. Pirotte MJ, Buckley BA, Lerhmann JF, Tanabe P. Development of a screening and brief intervention and referral for treatment for ED patients at risk for undiagnosed hypertension: a qualitative study. J Emerg Nurs 2014 Jan;40(1):e1-e9. [doi: 10.1016/j.jen.2012.05.004] [Medline: 23099012]
10. Moore T. How it works: Screening, Brief Intervention and Referral to Treatment (SBIRT). Delete: Delete; 2015.
11. Avis JL, Holt NL, Maximova K, van Mierlo T, Fournier R, Padwal R, et al. The development and refinement of an e-Health screening, brief intervention, and referral to treatment for parents to prevent childhood obesity in primary care. Telemed J E Health 2016 May;22(5):385-394. [doi: 10.1089/tmj.2015.0128] [Medline: 26451901]
12. Byrne JL, Wild TC, Maximova K, Browne NE, Holt NL, Cave AJ, et al. A brief eHealth tool delivered in primary care to help parents prevent childhood obesity: a randomized controlled trial. Pediatr Obes 2018 Nov;13(11):659-667. [doi: 10.1111/ijpo.12200] [Medline: 27863165]
13. Haff N, Patel MS, 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]

14. Mantzari E, Vogt F, Shemilt I, Wei Y, Higgins JP, 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]
15. Fernandes R, Chinn C, Li D, Halliday T, Frankland T, Ozaki R. Impact of financial incentives on health outcomes and costs of care among medicaid beneficiaries with diabetes in Hawai'I. Hawaii J Med Public Health 2019 Jan;78(1):19-25 [FREE Full text] [Medline: 30697471]
16. Look AHEAD Research Group, Wadden TA, West DS, Delahanty L, Jakicic J, Rejeski J, et al. The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. Obesity (Silver Spring) 2006 May;14(5):737-752 [FREE Full text] [doi: 10.1038/oby.2006.84] [Medline: 16855180]
17. Ryan DH, Espeland MA, Foster GD, Haffner SM, Hubbard VS, Johnson KC, Look AHEAD Research Group. Look AHEAD (Action for Health in Diabetes): design and methods for a clinical trial of weight loss for the prevention of cardiovascular disease in type 2 diabetes. Control Clin Trials 2003 Oct;24(5):610-628. [doi: 10.1016/s0197-2456(03)00064-3] [Medline: 14500058]
18. Look AHEAD Research Group, Pi-Sunyer X, Blackburn G, Brancati FL, Bray GA, Bright R, et al. Reduction in weight and cardiovascular disease risk factors in individuals with type 2 diabetes: one-year results of the look AHEAD trial. Diabetes Care 2007 Jun;30(6):1374-1383 [FREE Full text] [doi: 10.2337/dc07-0048] [Medline: 17363746]
19. Balamurugan A, Rivera M, Jack L, Allen K, Morris S. Barriers to diabetes self-management education programs in underserved rural Arkansas: implications for program evaluation. Prev Chronic Dis 2006 Jan;3(1):A15 [FREE Full text] [Medline: 16356368]
20. Lepard MG, Joseph AL, Agne AA, Cherrington AL. Diabetes self-management interventions for adults with type 2 diabetes living in rural areas: a systematic literature review. Curr Diab Rep 2015 Jun;15(6):608 [FREE Full text] [doi: 10.1007/s11892-015-0608-3] [Medline: 25948497]
21. Preechasuk L, Sriussadaporn P, Likitmaskul S. The obstacles to diabetes self-management education and support from healthcare professionals' perspectives: a nationwide survey. Diabetes Metab Syndr Obes 2019;12:717-727 [FREE Full text] [doi: 10.2147/DMSO.S195916] [Medline: 31190931]
22. Mehta S, Mocarski M, Wisniewski T, Gillespie K, Narayan KM, Lang K. Primary care physicians' utilization of type 2 diabetes screening guidelines and referrals to behavioral interventions: a survey-linked retrospective study. BMJ Open Diabetes Res Care 2017;5(1):e000406 [FREE Full text] [doi: 10.1136/bmjdrc-2017-000406] [Medline: 28878936]
23. Piccinino LJ, Devchand R, Gallivan J, Tuncer D, Nicols C, Siminerio LM. Insights from The National Diabetes Education Program National Diabetes Survey: opportunities for diabetes self-management education and support. Diabetes Spectr 2017 May;30(2):95-100 [FREE Full text] [doi: 10.2337/ds16-0056] [Medline: 28588375]
24. Rutledge SA, Masalovich S, Blacher RJ, Saunders MM. Diabetes Self-Management Education Programs in Nonmetropolitan Counties - United States, 2016. MMWR Surveill Summ 2017 Apr 28;66(10):1-6 [FREE Full text] [doi: 10.15585/mmwr.ss6610a1] [Medline: 28448482]
25. Wang J, Cai C, Padhye N, Orlander P, Zare M. A behavioral lifestyle intervention enhanced with multiple-behavior self-monitoring using mobile and connected tools for underserved individuals with type 2 diabetes and comorbid overweight or obesity: pilot comparative effectiveness trial. JMIR Mhealth Uhealth 2018 Apr 10;6(4):e92 [FREE Full text] [doi: 10.2196/mhealth.4478] [Medline: 29636320]
26. Brown-Guion SY, Youngerman SM, Hernandez-Tejada MA, Dismuke CE, Egede LE. Racial/ethnic, regional, and rural/urban differences in receipt of diabetes education. Diabetes Educ 2013;39(3):327-334. [doi: 10.1177/0145721713480002] [Medline: 23482514]
27. Lutfiyya MN, Patel YR, Steele JB, Tetteh BS, Chang L, Aguero C, et al. Are there disparities in diabetes care? A comparison of care received by US rural and non-rural adults with diabetes. Prim Health Care Res Dev 2009;10(4):320-331. [doi: 10.1017/s146342360999017x]
28. Health Resources and Services Administration. Defining Rural Populations URL: https://www.hrsa.gov/rural-health/about-us/definition/index.html [accessed 2020-03-06]
29. O'Brien T, Denham SA. Diabetes care and education in rural regions. Diabetes Educ 2008;34(2):334-347. [doi: 10.1177/0145721708316318] [Medline: 18375783]
30. Balamurugan A, Hall-Barrow J, Blevins MA, Brech D, Phillips M, Holley E, et al. A pilot study of diabetes education via telemedicine in a rural underserved community--opportunities and challenges: a continuous quality improvement process. Diabetes Educ 2009;35(1):147-154. [doi: 10.1177/0145721708326988] [Medline: 19244570]
31. Peng W, Yuan S, Holtz BE. Exploring the challenges and opportunities of health mobile apps for individuals with type 2 diabetes living in rural communities. Telemed J E Health 2016 Sep;22(9):733-738. [doi: 10.1089/tmj.2015.0180] [Medline: 26982017]
32. Lynch CP, Williams JS, JR K, GK R, E LE. Tablet-Aided BehavioraL intervention EffecT on Self-management skills (TABLETS) for Diabetes. Trials 2016 Mar 22;17:157 [FREE Full text] [doi: 10.1186/s13063-016-1243-2] [Medline: 27005766]

33. Onken LS, Carroll KM, Shoham V, Cuthbert BN, Riddle M. Reenvisioning clinical science: unifying the discipline to improve the public health. Clin Psychol Sci 2014 Jan 1;2(1):22-34 [FREE Full text] [doi: 10.1177/2167702613497932] [Medline: 25821658]
34. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: Key issues, process model, and system features. Commun Assoc Inf Syst 2009;24(1). [doi: 10.17705/1CAIS.02428]
35. Bang H, Edwards AM, Bomback AS, Ballantyne CM, Brillon D, Callahan MA, et al. Development and validation of a patient self-assessment score for diabetes risk. Ann Intern Med 2009 Dec 1;151(11):775-783 [FREE Full text] [doi: 10.7326/0003-4819-151-11-200912010-00005] [Medline: 19949143]
36. Miller W, Rollnick S. Motivational Interviewing: Preparing People for Change. Second Edition. New York: Guilford Press; 2002.
37. American Diabetes Association. Abridged for primary care providers. Clin Diabetes 2017 Jan;35(1):5-26 [FREE Full text] [doi: 10.2337/cd16-0067] [Medline: 28144042]
38. Parra-Medina D, Wilcox S, Wilson DK, Addy CL, Felton G, Poston MB. Heart Healthy and Ethnically Relevant (HHER) Lifestyle trial for improving diet and physical activity in underserved African American women. Contemp Clin Trials 2010 Jan;31(1):92-104 [FREE Full text] [doi: 10.1016/j.cct.2009.09.006] [Medline: 19781665]
39. Clark CM, Fradkin JE, Hiss RG, Lorenz RA, Vinicor F, Warren-Boulton E. The National Diabetes Education Program, changing the way diabetes is treated: comprehensive diabetes care. Diabetes Care 2001 Apr;24(4):617-618. [doi: 10.2337/diacare.24.4.617] [Medline: 11315818]
40. Funnell MM, Brown TL, Childs BP, Haas LB, Hosey GM, Jensen B, et al. National standards for diabetes self-management education. Diabetes Care 2009;32(Supplement_1):S87-S94. [doi: 10.2337/dc09-s087]
41. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
42. Sheeran P, Maki A, Montanaro E, Avishai-Yitshak A, Bryan A, Klein WM, et al. The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: A meta-analysis. Health Psychol 2016 Nov;35(11):1178-1188. [doi: 10.1037/hea0000387] [Medline: 27280365]
43. McConnon Á, Kirk SF, Ransley JK. Process evaluation of an internet-based resource for weight control: use and views of an obese sample. J Nutr Educ Behav 2009;41(4):261-267. [doi: 10.1016/j.jneb.2008.07.008] [Medline: 19508931]
44. Strom JL, Lynch CP, Egede LE. Rural/urban variations in diabetes self-care and quality of care in a national sample of US adults with diabetes. Diabetes Educ 2011;37(2):254-262. [doi: 10.1177/0145721710394875] [Medline: 21289298]
45. Young L, Barnason S. Review strategies to recruit and retain rural patient participating self-management behavioral trials. Online J Rural Res Policy 2015;10(2):1-12 [FREE Full text] [doi: 10.4148/1936-0487.1070] [Medline: 28580049]
46. Young L, Montgomery M, Barnason S, Schmidt C, Do V. A conceptual framework for barriers to the recruitment and retention of rural CVD participants in behavior intervention trials. GSTF J Nurs Health Care 2015 Aug;2(2):19-24 [FREE Full text] [doi: 10.5176/2010-4804_2.2.71] [Medline: 28580381]
47. Monitoring and Evaluating Digital Health Interventions: A Practical Guide to Conducting Research and Assessment. Geneva: World Health Organization; 2016.

Abbreviations

ADA: American Diabetes Association
CHW: community health worker
mHealth: mobile health
MI: motivational interviewing
RCT: randomized controlled trial
SBIRT: Screening, Brief Intervention, and Referral to Treatment
T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 15.10.19; peer-reviewed by E Bellei, D Biduski, J Segel; comments to author 11.11.19; revised version received 23.12.19; accepted 26.01.20; published 24.03.20.

Please cite as:
Yin Z, Lesser J, Paiva KA, Zapata Jr J, Moreno-Vasquez A, Grigsby TJ, Ryan-Pettes SR, Parra-Medina D, Estrada V, Li S, Wang J Using Mobile Health Tools to Engage Rural Underserved Individuals in a Diabetes Education Program in South Texas: Feasibility Study
JMIR Mhealth Uhealth 2020;8(3):e16683
URL: http://mhealth.jmir.org/2020/3/e16683/
doi:10.2196/16683
PMID:32207694

Original Paper

A Mobile Phone Intervention to Improve Obesity-Related Health Behaviors of Adolescents Across Europe: Iterative Co-Design and Feasibility Study

Anne Martin[1*], BSc, MSc, PhD[‡]; Maurizio Caon[2*], BSc, MSc, PhD; Fulvio Adorni[3], MSc; Giuseppe Andreoni[4], BSc, PhD; Antonio Ascolese[5], PhD; Sarah Atkinson[6], BSc, MSc, PhD, CErgHF; Kim Bul[7], BSc, MSc, PhD; Carme Carrion[8], PhD; Conxa Castell[9], MD, PhD; Valentina Ciociola[3], MSc; Laura Condon[10], PhD; Mireia Espallargues[11,12], MPH, MD, PhD; Janet Hanley[13], BSc, MPhil, PhD; Nithiya Jesuthasan[3], MSc; Claudio L Lafortuna[14], MD; Alexandra Lang[6], BSc, PhD; Federica Prinelli[3], PhD; Elisa Puidomenech Puig[11,12], BSc, MPH; Sarah A Tabozzi[14], MSc; Brian McKinstry[15], MD, FRCPE, FRCGP

[1]UK Medical Research Council / Chief Scientist Office Social and Public Health Sciences Unit, University of Glasgow, Glasgow, United Kingdom

[2]Haute Ecole de Gestion and Haute Ecole d'Ingénierie et d'Architecture, University of Applied Sciences and Arts of Western Switzerland, Fribourg, Switzerland

[3]Institute of Biomedical Technologies, National Research Council, Milan, Italy

[4]Design Department, Politecnico di Milano, Milan, Italy

[5]Imaginary srl, Milan, Italy

[6]Human Factors Research Group, University of Nottingham, Nottingham, United Kingdom

[7]Centre for Innovative Research Across the Life Course, Coventry University, Coventry, United Kingdom

[8]eHealth Lab Research eGroup, Universitat Oberta de Catalunya, Barcelona, Spain

[9]Public Health Agency of Catalonia, Barcelona, Spain

[10]Division of Primary Care, University of Nottingham, Nottingham, United Kingdom

[11]Catalan Agency for Health Information, Assessment and Quality, Barcelona, Spain

[12]Health Services Research on Chronic Patients Network, Barcelona, Spain

[13]School of Health and Social Care, Edinburgh Napier University, Edinburgh, United Kingdom

[14]Institute of Molecular Bioimaging and Physiology, National Research Council, Milan, Italy

[15]Usher Institute for Population Health Sciences, University of Edinburgh, Edinburgh, United Kingdom

[‡]Written on behalf of PEGASO Fit For Future consortium

[*]these authors contributed equally

Corresponding Author:
Anne Martin, BSc, MSc, PhD
UK Medical Research Council / Chief Scientist Office Social and Public Health Sciences Unit
University of Glasgow
200 Renfield Street
Glasgow, G2 3AX
United Kingdom
Phone: 44 141353 ext 7603
Email: Anne.Martin@glasgow.ac.uk

Abstract

Background: Promotion of physical activity, healthy eating, adequate sleep, and reduced sedentary behavior in adolescents is a major priority globally given the current increase in population health challenges of noncommunicable diseases and risk factors such as obesity. Adolescents are highly engaged with mobile technology, but the challenge is to engage them with mobile health (mHealth) technology. Recent innovations in mobile technology provide opportunities to promote a healthy lifestyle in adolescents. An increasingly utilized approach to facilitate increased engagement with mHealth technology is to involve potential users in the creation of the technology.

Objective: This study aimed to describe the process of and findings from co-designing and prototyping components of the PEGASO Fit for Future (F4F) mHealth intervention for adolescents from different cultural backgrounds.

Methods: A total of 74 adolescents aged 13 to 16 years from Spain, Italy, and the United Kingdom participated in the co-design of the PEGASO F4F technology. In 3 iterative cycles over 12 months, participants were involved in the co-design, refinement, and feasibility testing of a system consisting of diverse mobile apps with a variety of functions and facilities to encourage healthy weight–promoting behaviors. In the first iteration, participants attended a single workshop session and were presented with mock-ups or early-version prototypes of different apps for user requirements assessment and review. During the second iteration, prototypes of all apps were tested by participants for 1 week at home or school. In the third iteration, further developed prototypes were tested for 2 weeks. Participants' user experience feedback and development ideas were collected through focus groups and completion of questionnaires.

Results: For the PEGASO F4F technology to be motivating and engaging, participants suggested that it should (1) allow personalization of the interface, (2) have age-appropriate and easy-to-understand language (of icons, labels, instructions, and notifications), (3) provide easily accessible tutorials on how to use the app or navigate through a game, (4) present a clear purpose and end goal, (5) have an appealing and self-explanatory reward system, (6) offer variation in gamified activities within apps and the serious game, and (7) allow to seek peer support and connect with peers for competitive activities within the technology.

Conclusions: Incorporating adolescents' preferences, the PEGASO F4F technology combines the functions of a self-monitoring, entertainment, advisory, and social support tool. This was the first study demonstrating that it is possible to develop a complex mobile phone-based technological system applying the principles of co-design to mHealth technology with adolescents across 3 countries. The findings from this study informed the development of an mHealth system for healthy weight promotion to be tested in a controlled multinational pilot trial.

(JMIR Mhealth Uhealth 2020;8(3):e14118) doi:10.2196/14118

KEYWORDS

health behavior; obesity; co-design; mHealth; mobile app; mobile phone; adolescents; youth; focus groups

Introduction

Background

Innovations in mobile technology for health (mobile health [mHealth]) provide a novel opportunity for combating child and youth overweight and obesity by promoting behavior change for increased physical activity, reduced sitting time, and healthy eating habits [1]. Recent developments of mobile technology provide personalized, continuous, and context-aware feedback on behavior, which is suggested to be a critical component for health behavior change and potentially increases adherence [2-4]. Integrating game design elements in existing intervention approaches (so-called gamification) and using educational games (ie, serious games) to improve health-related behavioral outcomes appear to be effective in young people [5,6]. However, in the context of promoting a healthy lifestyle among young people, more research efforts need to be directed to successfully promoting adequate engagement with these apps as high levels of disengagement with mHealth technology have been reported [7]. This long-term disengagement will result in nonadoption and eventual abandonment of the technology by intended users [8]. Although young people are becoming more interested in using commercial mobile apps for their health [9], engagement with evidence-based mHealth interventions is low. For example, adolescents participating in a 12-week mobile phone-assisted healthy weight intervention monitored their meals on only 16.6% of available days and tracked at least 30 min of physical activity on only 4.6% of available days [10]. In an 11-week randomized controlled trial for testing a mindfulness app for improving weight-related behaviors, only 14% (5/34) of young adults accessed the content regularly, 61% (21/34) used the app periodically and 23% (8/34) did not engage with the app [11]. This literature exemplifies the challenges and complexity of developing and implementing a successful mHealth intervention. The need and potential benefit of the intervention may be clear, but the initial uptake and long-term use will be contingent on the technology's ability to eliminate barriers to engagement via integrating and adapting to real-world complexity [8].

One way of tackling disengagement is the user-centered design (UCD) approach, which puts the user, with their needs, desires, and limitations at the center of the design process. The value of this approach is now well recognized, and as such, to improve engagement and uptake of mHealth by young people, they are increasingly being involved in the creation of these technologies. Co-design refers to the collective "creativity of designers and people not trained in design working together in the design development process" [12]. Co-design is an approach that is a subdiscipline of participatory design and can be integrated into a UCD process, ensuring that end users are provided with opportunities to express what they want from a product or service. This can be achieved using qualitative and design methodologies that are tailored to the specific user group in question. Co-design provides designers with the right tools to elicit reliable user requirements and subsequently develop systems that better meet user needs and improve user experience [13]. The importance of implementing this approach and involving multiple stakeholders in the development of mHealth interventions has been demonstrated in the literature supporting mHealth for nutrition management and in its specific application in co-design with young people [14-16].

Objectives

Recognizing the potential of mHealth as an efficient tool for health behavior change in young people and the need for

adequate engagement with the technology, the PEGASO Fit for Future (F4F) project aimed to develop a multidimensional interdisciplinary system placing the target population in the center of development process. The PEGASO F4F system exploited sophisticated game mechanics aiming to motivate behavioral changes toward a healthier lifestyle and prevention of youth obesity. The 3 main features of the PEGASO F4F system were: (1) individual and environmental monitoring, including wearable sensors, mobile phones, and multimedia diaries for the acquisition of physical, physiological, and behavioral attributes of participants; (2) feedback to the user, presenting personalized healthy options for alternative lifestyles; and (3) social connectivity and engagement, encouraging involvement in social network and sharing experiences [17]. The PEGASO F4F system was intended to function on its own and was not part of a wider nondigital healthy weight–promotion initiative or intervention. The entire PEGASO F4F system comprised 8 connected technological components; however, the focus of this paper was to report on the development of 3 system components. The development and content of the remaining components were described elsewhere [18]. All system components were intended to function in conjunction with each other rather than as a standalone application.

The development of the PEGASO F4F system was aligned to the Integrate, Design, Assess, and Share framework for the development of digital technology for health behavior change [19]. The framework consists of 10 phases: (1) empathize with target users, (2) specify target behavior, (3) ground in behavioral theory, (4) ideate implementation strategies, (5) prototype potential products, (6) gather user feedback, (7) build a minimum viable product, (8) pilot test to assess potential efficacy and usability, (9) evaluate efficacy in a randomized controlled trial, and (10) share intervention and findings. Phases 1 to 4 of the PEGASO F4F project were described elsewhere [20,21].

Phases 5 to 7 are the focus of this paper. Therefore, the aim of this study was to describe the process and findings of co-designing and prototyping potential components of the PEGASO F4F mHealth intervention and building a minimum viable product before testing the feasibility and piloting the PEGASO F4F system in a controlled multicenter intervention study (phase 8). Phases 1 to 8 were funded by the European Commission under the Seventh Framework Programme between 2014 and 2017. The research presented in this paper took place between 2015 and 2016.

Methods

Study Design

PEGASO F4F was a European multicenter project. Co-design and development of the technology took place in 3 iterative stages in the United Kingdom (England and Scotland), Italy (Lombardia), and Spain (Catalonia). Table 1 summarizes the specific objectives for the PEGASO F4F apps and serious game. The objectives of subsequent iterations were influenced by the findings and the co-design progress of the previous iteration.

Table 1. Objectives for each co-design phase.

System component	First iteration	Second iteration	Third iteration
Apps	Assessment of usability and satisfaction, comparison of 2 different prototypes with different interaction layers	Assessment of usability and satisfaction; testing of automatic functionalities, user data entry functionalities	Assessment of usability, intuitiveness, and satisfaction; testing of functionalities
Serious game	Assessment of user needs and acceptance of the visual and graphical content	Assessment of user needs, usability, and acceptance of the game mechanics	Assessment of user needs, usability, and acceptance of the game mechanics; enjoyment, intuitiveness, and playtime

Participant Recruitment and Study Setting

Male and female adolescents in Spain, Italy, England, and Scotland between the ages of 13 and 16 years were invited through their schools to participate in this study. Convenience sampling was employed for selecting schools. However, schools located in both urban and rural areas and from areas of high and lower deprivation were recruited. Participants were expected to be technologically savvy, that is, with experiences, interest, and ability in using technological devices. This information was based on participants' self-belief about their interest in and experience with technology. Adolescents whose English language skills did not allow for understanding of the study material were excluded from this study as, for practical reasons, all mock-up and prototype material was available in English language only. In 3 of the 4 study sites, different adolescents were recruited for each iterative stage to ensure the results are generalizable to the general adolescent population. Written study information was provided and written informed consent obtained from the participants directly (Scotland) or from their parents (England, Spain, and Italy). All 3 iterative stages of the technology co-design process were approved by the research ethics committees from the responsible institution in each country: England (Faculty of Engineering Ethics Committee, The University of Nottingham), Scotland (National Health Service South East Scotland Research Ethics Committee 16/SS/0163), Spain (Ethics Committee of Institut Universitari d'Investigació en Atenció Primària Jordi Gol), and Italy (Istituto di Ricovero e Cura a Carattere Scientifico Policlinico of Milan Ethics Committee).

Concepts of the PEGASO Fit for Future System Components

The development of the PEGASO F4F system was guided by 4 behavioral and theoretical frameworks to raise participants' awareness about their health behavior and tailor adequate actions for behavior change, namely the Behavior Change Wheel [22], positive psychology [23], self-determination theory [24], and nudging theory [25]. In all, 2 different persuasive mechanisms—gamification and serious gaming—were proposed

to allow the participant to engage with the technology while integrating strategies for behavior change [26].

Initial design ideas were derived from the literature and 3 focus groups with 27 young people aged 13 to 15 years [18]. The focus groups explored how young people interact with technology and their preferences for doing so in relation to promoting physical activity and healthy eating, thus informing the prototypes and mock-ups developed for iteration 1 of this study [18].

The entire PEGASO F4F system comprised 8 technological components; however, the focus of this paper was to report on the development of the Companion app, the serious game, and the eDiary app.

Description of PEGASO Fit for Future System Components

Companion App

The objective of the Companion app was to encompass all other PEGASO F4F apps providing a seamless and unique experience to the user. The Companion app was intended to be the app through which the user could engage with all other PEGASO F4F system components. Therefore, users would open one app that in turn contained a suite of functionalities, some of which would be complex apps such as the eDiary and serious game. In addition, the Companion app was intended to make an active lifestyle more attractive, to support users to eat healthier, and to connect users with their peers. To achieve this, notification messages were proposed as a mechanism to facilitate the 3 main functions of the app [27]. The notifications had the potential to be motivational, educational, or acted as reminders to engage with other components of the PEGASO F4F system. For example, notifications reminded the user to enter food items to select a challenge or to nominate a friend for peer support. The challenge component of the Companion app allowed the users to create and accept different types of challenges relating to physical activity and healthy eating. This app offered the opportunity for participants to challenge themselves and other users in a competitive or collaborative way. Previous literature indicated that collaboration and competition can both be beneficial for promoting motivation [28]. The Companion App also included a reward section dedicated to the badges enabling users to unlock new app content and achieve specific goals by using 1 or multiple PEGASO F4F apps and performing the suggested activities. Badges performed multiple functions in congratulating the user on achievements and promoting the social experience of PEGASO F4F through the sharing of badges. This gamification element was intended to generate fun while motivating users to perform specific tasks as part of their interventions.

Serious Game

Using a postapocalyptic zombie narrative, the serious game aimed to be the motivational component of PEGASO F4F. It needed to be immersive and engaging for the player, while utilizing the PEGASO F4F system to capture information about lifestyle and encourage maintenance or change of health behaviors. In all, 2 central behavioral mechanisms were envisioned: (1) utilizing behavioral theories of self-determination and nudging, an energy bar indicated the change in energy available for playing the game—the player's actions in the game reduce energy, whereas achieving real-life behavioral goals such as increased physical activity or improved diet would increase the available energy to play the game, and (2) the game implements *research* mechanics that require the player to apply and develop their nutritional knowledge of various food sources by playing minigames. Minigames were short games embedded within the overall game environment creating opportunities to gain knowledge on food and develop new nutrition-related skills. In the final version of PEGASO F4F, progress within the serious game would be contingent on the physical activity levels of the participant, as measured by the wearable sensors integrated into the system.

eDiary App

The objective of the eDiary app was to enable the user to enter information about their usual diet and to provide feedback about healthy eating habits. This app could also provide suggestions to the user on how to improve their dietary habits (eg, increasing or reducing certain food groups and recipes). Informed by literature, researchers with specialist expertise in nutrition, proposed to focus the content of the eDiary app on 6 target behaviors: increased consumption of fruit, increased consumption of vegetables, reduced intake of fast food, reduced intake of sugar-sweetened beverages, healthy snacking, and reduced breakfast skipping. Those behaviors have been selected because of their relevance in obesity prevention and their ability to be modified and measured [29]. The eDiary app allowed the user to select behavioral goals for one of the 6 target behaviors. Users also received automated feedback on the diversity (ie, degree of variation of the diet) and balance (ie, adequacy of food servings relative to recommendations) of the consumed food.

Co-Design and Technology Development Procedures

The activities, objectives, and duration of each co-design iteration varied depending on the development stage of each PEGASO F4F component. Table 2 summarizes the characteristics of the 3 co-design iterations.

Table 2. Components and development stages of the PEGASO Fit for Future Companion app, eDiary app, and serious game.

Stages of system development	First iteration	Second iteration	Third iteration
Mock-up	Companion app and serious game	—[a]	—
Prototype 1.0	eDiary app	Companion app and serious game	—
Prototype 2.0	—	eDiary app	Companion app and serious game
Prototype 3.0	—	—	eDiary app
Integration status	Separate, not integrated system components	Separate, not integrated system components	Partially integrated system components

[a]Progress in development of system components leads to empty cells.

To allow adolescents from a wide range of socioeconomic backgrounds to participate in this study, mobile phone handsets with Android OS version 5.0 (and higher) were provided to participants for the duration of the study.

In the first iteration, participants attended a single workshop session in which they were shown screenshots of the serious game and the Companion app. Participants were presented with the first prototype of the eDiary app, and they were asked to carry out 2 tasks to assess the usability of the app and examine how young people engaged with the early prototype. Task 1 was to modify the entry of a meal, and task 2 was to ask for dietary feedback based on the inserted example meals (example meal plans were provided by the researchers).

During the second iteration, prototypes of both apps and the serious game were trialed on a daily basis for 1 week in the home and school context. The data from the wearable activity monitors were not yet integrated into the serious game because of parallel development work streams. This meant that the serious game was initially tested with in-game progression being dependent on time rather than the physical activity levels of the participant. Future prototype testing (outside the scope of this study) would see the gatekeeping mechanism to in-games progression regarding *energy levels* and accessing new game content being driven by the participants activity levels as measured by the wearable sensors. For this iteration, energy within the game increased slowly over time. This served the purpose of testing if adolescents could be encouraged to engage in short daily game sessions rather than prolonged single-game sessions.

In the third and final stage of the co-design phase, the eDiary and serious game apps were integrated within the Companion app (Table 2). For this iteration, progression in the serious game was based on the player's physical activity captured by the accelerometer in the mobile phone. Participants of the third co-design iteration tested the PEGASO F4F components for 2 weeks.

Data Collection and Analysis

Table 3 summarizes the data collection methods for gathering participants' feedback for each tested PEGASO F4F system component. A mixed method co-design approach was chosen to collect as much feedback in breadth and depth as possible to inform the development of the PEGASO F4F system. In addition, 2 experienced researchers were present during all the workshop and focus group sessions. One researcher was moderating the session and another was allocated to make notes, keep track of the time, and make observations. Although the system components were only available in English, all workshop activities and questionnaires were conducted in the local language.

Table 3. Data collection methods used for each PEGASO component at each iteration.

Data collection method	First iteration	Second iteration	Third iteration
Focus groups	eDiary app and serious game	eDiary app, Companion app, and serious game	eDiary app, Companion app, and serious game
System Usability Scale	—[a]	eDiary app and serious game	eDiary app, Companion app, and serious game
Other	App design questionnaire and character design worksheet	Brief Use of App questionnaire and survey on game use	Survey on game use

[a]Not applicable.

Design Activities

A character sketchbook with many different proposals for the avatars, icons, and other design elements was provided by developers as a resource for participants to review and develop the Companion and eDiary apps (Multimedia Appendix 1). Vote counting determined the most preferred design.

Focus Groups

Focus groups were conducted at all 4 test sites in the local languages and at school and were facilitated by local postdoctoral researchers with a background in behavioral sciences or human-computer interaction (HCI). Facilitators were both male and female in Scotland, Spain, and Italy and all female in England and were not directly involved in the development of the PEGASO F4F software. The focus group topic guides

for all 3 iterations can be found in Multimedia Appendix 2 (English-language version only). Focus groups lasted between 60 and 90 min in each iteration, and they were voice or video recorded, transcribed, and thematically analyzed at each study site. Weight of opinion (in terms of number of participants agreeing) was used to make decisions regarding the design, look, and feel or the apps and serious game. Similarities and differences in participants' views between study sites were identified using thematic analysis, whereby similar feedback across the sites was considered for incorporation in the technology development.

Questionnaires

For the development of the Companion app, eDiary app, and serious game, participants were asked to complete 5 questionnaires over the course of the 3 iterative circles (Table 3).

Participants were asked to complete the System Usability Scale (SUS [30]) to determine participants' subjective rating of the usability of the apps and the serious game. The SUS index was calculated using the standard conversion method [30,31], with scores above 70 indicating acceptable usability, scores between 50 and 70 suggesting marginal usability, and scores below 50 reflecting poor or unacceptable usability [31]. The questionnaire data were entered in to Microsoft Excel and IBM SPSS software (version 21), descriptive summary statistics calculated.

The Brief Use of App Questionnaire (Multimedia Appendix 3) was developed by the research team and included closed and open-ended questions. It consisted of 7 questions relating to the Companion App, of which 3 questions aimed to obtain information on participants' preferences about the *challenges*, and 4 questions related to the use of the eDiary app.

The Survey on Game Use was also researcher derived and consisted of 4 questions to determine (1) how long participants played the game during a week, (2) when and where they played the game, (3) whether and with whom they talked about the game, and (4) suggestions for the next version of the serious game.

Results

Participant Characteristics

A total of 74 adolescents aged between 13 and 16 years participated in the co-design of the PEGASO F4F technology. The number of participants varied between the 4 study sites and between the 3 iterations (Table 4).

Table 4. Participant characteristics for each study site and co-design iteration.

Participant characteristics	First iteration	Second iteration	Third iteration
Participants, n			
Spain	8-10[a]	9	10
England	16[b]	10[b]	10[b]
Italy	9-10[a]	7-12[c,d]	10[d,e]
Scotland	11	8-9	5
Sex (males:females)			
Spain	7:3, 2:6	5:4	5:5
England	6:9	7:3	3:7
Italy	6:4, 2:7	6:6	4:4
Scotland	6:5	5:4	3:2

[a]10 participants tested the serious game and 8-9 tested the companion and eDiary apps.

[b]1 participant took part in all 3 iterations, 1 participant took part in the first and second iterations, and 4 participants took part in the second and third iterations.

[c]The number of completed questionnaires varied for different apps.

[d]The same participants took part in the second and third iterations.

[e]10 participants tested the apps.

Participant Feedback and Design Ideas

Companion App

Feedback from the first iteration suggested that the concept of the Companion app was perceived positively and was well accepted by all participants. Participants understood the importance of a central personalized platform to introduce them to the wider PEGASO F4F system. Participants suggested alternative design options for two out of six icons of the Companion app, provided feedback with regards to a PEGASO mascot and a Companion app avatar (Table 5). All participants responded positively to the use of emoticons but suggested the adoption of existing libraries (eg, WhatsApp). The concept of data sharing was an important topic for which participants had clear views. Participants suggested that sharing of the full name and a profile was appropriate at all times. However, data on eating habits should not be shared with peers at any point. As a default, this was agreed, with the addition of modifiable settings that were controlled by the user to share this information but only when and with whom they wanted.

Participants who tested the next version of the Companion app in the third iteration of the study (Figure 1), correctly understood the key functions and reported that it was easy to use. They found it visually attractive, helpful, and usable. The average SUS scores indicated marginally acceptable usability of the app in Italy 63.0 (SD 15.1), Scotland 60.0 (SD 17.3), and England 63.3 (SD 36.2) and poor acceptability in Spain 43.7 (SD 4.2). Participants reported that the Companion app was useful as a single app portal for all PEGASO F4F components. However, some of the participants accessed other PEGASO F4F apps bypassing the Companion app. Other participants did not realize that it was possible to collect points and rewards by using the Companion app. Participants of the third iteration liked the idea of being able to set progressive challenges and their own goals, broadening the variety of challenges and sorting them by areas such as *sports* or *eating*. Checking personal progress was considered as motivating for engagement in a healthy lifestyle. They also suggested getting real-life rewards for every challenge met, for example, using social media for setting challenges, where invitations to sport events could be sent and shared.

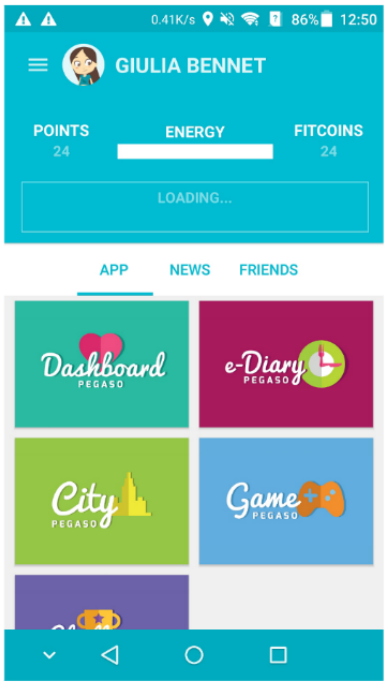
When using the first prototype of the Companion app in the second iteration, participants experienced difficulties in logging on and in getting the app running. Once it was in operation, participants used the app several times a week and found it easy to use. Participants liked receiving notification messages but expressed concerns about the number and timing of incoming messages (Table 5). For example, adding dietary information to the app while being in class. Participants liked the messages to be short and cheerful in tone, aiming to motivate and provide reminders. Regarding the feedback on health behavior change, most of the participants said that being rewarded with medals that were visible to friends would be seen in a positive way as being motivating to them. The iterative co-design of the visual elements of the interface meant that participants liked the aesthetics of the Companion app; they found the layout was easy to navigate and was easy to read. Most participants found that the amount of information and arrangement on the screens was logical, with clear sequencing of the screens and a predictable ordering of what the next screen advancement would be, suggesting that the early app development complied with the HCI usability principles of visibility of system status and match between system and real world and positive aesthetic experience [32]. Participants liked the idea of being able to set challenges for their friends, compete with each other, and share the results.

Table 5. Participants' co-design output and implemented solution for each PEGASO Fit for Future system component.

Component and participant feedback	Implemented solutions
Companion app	
Not understanding the meaning of the Health Square app icon and Challenge icon; alternative icon design suggested.	The Challenge icon was changed.
Liked the idea of having a mascot and selected a favorite design.	A Mascot was added based on the design selected by the majority of participants.
Having a customizable avatar.	Avatars were added based on selected design.
Not wanting to share data around eating habits.	Sharing of information was limited with control on whether sharing and when.
Some challenges were difficult, and some were easy; suggested to include incremental challenges.	Incremental challenges implemented.
Participants suggested to include a leader board as they would be more motivated to engage in health behaviors when seeing their friends' performance.	Leader board was added.
Did not like to receive too many notifications or messages.	The number of notifications was fixed to 2/4 messages per day.
Frustration when receiving notification when it is not possible to respond; suggested to receive notifications in after-school hours.	Notifications timed to be released to later in the day (after school).
Needing a tutorial or guidance for use of the app.	Tutorial messages implemented.
Participants did not realize that it was possible to collect points and rewards.	Added a tutorial message and an explanation in the *Points* section.
Bypassing the companion app to use other PEGASO F4F apps.	Bypassing the Companion app to access other PEGASO F4F apps will not be possible as soon as all remaining PEGASO F4F system services are integrated.
Serious game	
Did not understand the story narrative of the story; suggested to add a tutorial on what the game is about.	A video introduction and tutorial were created.
The game world is too easy to explore (suggested to provide further dynamism and complexity; specific design ideas for environments or locations provided).	New locations were added. The number of locations increased from 1 to 4.
Repeating the same minigame was boring.	2 additional minigames were included (scavenging and research).
No ranks and competitive elements; suggested to include a leader board.	Leader board was implemented.
Some participants noted they were progressing to higher levels, but they did not notice clear differences between the levels; suggested to add new type of zombies, environments, and abilities (specific ideas provided) with each level.	The possibility to unlock (in different levels) new zombies, new environments, and new abilities were added.
Suggested to add audio features and sounds.	Music and audio effects were added to increase engagement.
Feeling that actions in the game had no concrete relationship to the Companion app; suggested to gain coins that can be used in other parts of the PEGASO F4F system.	Fit coins to be used in other sections in the Companion app were introduced.
eDiary app	
It was difficult for adolescents to understand the servings and food included in a group; suggested to include a tutorial.	A help function in form of a question mark icon was added to provide guidance for the food input.
Preference of symbolic food icon over food images.	Implemented the icon interface instead of the one based on food images.
Some icons were difficult to be interpreted; alternative images were suggested.	Some food icons (eg, fried food, soft drink, Asian food, and snacks) were redesigned following participants' suggestions.
Selected a favorite design for the graphical feedback on diet.	Design preference implemented.
Not understanding the meaning of "equilibrium."	"Equilibrium" changed to "balance."
Not understanding the meaning of the "diversity" and "balance."	Explanation of the indexes added.
Wanting suggestions on how to improve their diet.	Generic recommendations were shown in the News Stream and personalized recommendations were implemented in the eDiary.

Component and participant feedback	Implemented solutions
More instructions were needed.	Tutorial cards as instructions and improved help text added.
Participants proposed to increase the number of food groups and suggested alternative food groups.	Food groups were restructured adding new icons. For example, the food groups *fast* food and *snacks* were represented in more detailed food icons for "Burger," "Oriental," "Chicken," "Pizza," "Salty," "Bakery," "Sweets & Dessert."
Preference of receiving advice on which foods should be eaten if a person has food allergies or intolerances.	This was not implemented because of lack of resources.

Figure 1. Screenshot of the PEGASO Fit For Future Companion app presented at the third iteration.



Serious Game

Based on the feedback from the first iteration, participants were positive about the main idea of the game, mainly because of the novelty of the topic and the setting of the game. They liked the idea that they needed to move and to spend energy in real life to gain more energy in the game, a concept which was only tested in the third iteration (see Multimedia Appendix 4). Participants felt that the game was suitable for a mobile phone. On the basis of the mock-up presentation, it became evident that some aesthetic elements should be improved, but the icons and the images were understandable in general. There was a lack of consensus regarding the graphics of the game: some participants preferred the cartoon style of game characters and others the realistic and more sophisticated look of characters. Participants suggested that they could be more engaged with the game if they could interact with other players. They said that they would prefer to customize their avatar (eg, gender, physical appearance, armor, and weapons), to choose their attributes, and to increase skills during the game.

Most of the participants played the serious game between 1 and 3 hours during the weeklong second iteration of this study. Participants played the game in a variety of locations and most of them at home, indicating that the game was usable and acceptable in the daily context of young people's lives. The suggested features that they wanted to see in the next version of the game included (1) more worlds to explore, (2) more types of zombies and characters, and (3) rewards for healthy real-world actions. Some participants asked for tutorials and help during the game because it was unclear how the game worked and what the aim of the game was. Regarding the game play mechanics, most of the participants felt the daytime activity of finding foods was boring. They suggested adding missions or other game features, such as exploring the inside of the buildings or driving cars. Participants asked for a variety of weapons, more variety of in-game characters, ways to fight zombies, and ways to increase their character abilities. Some participants did not understand how to play the *minigame* and what to do with the recipe cards. The participants felt that they needed more instructions on how to do different aspects of the serious game. The usability of the game using the SUS scale was rated as marginally usable by participants in Spain (mean 67.8, SD 18.2), Italy (mean 62.8, SD 10.8), and Scotland (mean 66.8, SD 17.1). However, participants in England rated the game as having poor usability (mean 20.3, SD 12.0).

Most of the participants of the third iteration of this study liked the concept of the game and said that the controls were easy to use. However, they found that the controls reacted too slowly,

and the game was considered repetitive with a lack of progression or a clear end goal. Some participants noted progressing to higher levels in the game, but they did not notice a clear difference between levels. The need of a tutorial within the game was expressed again to understand the aim of the game and *minigames* and how to play them. The written manual provided in response to feedback of the second iteration did not offer adequate assistance to participants of the third iteration. Participants also suggested including different scenarios, more activities, and a clear goal or different *minigoals* that can be unlocked after gaining points. In addition, participants suggested having different settings that could change as the player progressed through the game as a way of reducing repetitiveness. The revised minigames were perceived as visually attractive and usable. However, the participants did not welcome repeating the same *minigame*, and they suggested having additional minigames. Perceptions about the usability of the game using the SUS were mixed across countries. The serious game reached acceptable usability in England (mean 75.0, SD 6.4), marginal usability in Italy (mean 54.3, SD 12.2) and Scotland (mean 57.5, SD 21.1), and poor usability in Spain (mean 45.3, SD 14.1).

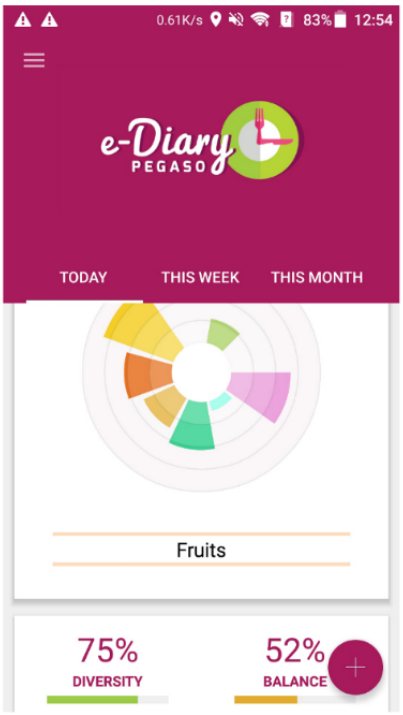
eDiary App

The idea of the eDiary app was welcomed by females more than males after testing the first prototype of the app in the first iteration of this study. Participants expressed that they would like a tool that is able to collect information about nutritional contents of foods and beverages, to indicate daily dietary balance, and to advice on which foods should be eaten if a person has food allergies or intolerances. The initial food icon designs were well understood by most participants; however, where there was confusion or low satisfaction with the proposed designs, the co-design process enabled participants to offer changes and suggested icons to improve clarity. Participants also expressed their preference of the design of the graphical feedback display. Participants said that they preferred a user interface that was more symbolic rather than displaying photos of food items. Several participants proposed to increase the number of food groups and to reduce the number of foods included in each food group. Participants liked the idea of receiving feedback on how balanced their daily diet was.

During the second iteration, all participants used the eDiary app, and generally, females were more willing to use this app regularly than males. All the participants noted that the food groups were too limited and sometimes did not fit with the information that participants wanted to enter and felt that the food groups should be better explained. Participants noted that entering information on food consumed was difficult at times, leading to frequent mistakes that were difficult to erase; thereby, indicating that the system did not enable users to recognize, diagnose, and recover from errors as is recommended in the HCI literature [32]. Moreover, participants did not like the negative feedback when the food intake was not healthy. They suggested the app should tell them how to improve their eating habits. Mean SUS scores indicated marginal usability across all countries: Spain (mean 58.3, SD 18.6), Italy (mean 70.0, SD 14.6), Scotland (mean 69.1, SD 16.4), and England (mean 59.3, SD 7.6).

Participants who used the revised prototype version of the eDiary app in the third iteration (Figure 2) found the app easy to understand and aesthetically pleasing. All the participants had used eDiary and agreed that they liked it. This feedback aligned with the improved usability ratings, with average SUS scores of 61.2 (SD 7.0), 75.0 (SD 16.8), 71.5 (SD 13.5), and 76.8 (SD 16.7) in Spain, Italy, Scotland, and England, respectively. Participants liked to get feedback about their food intake and the visualized feedback in form of a pie chart (Figure 2) helped them to increase their perceived awareness on their diet. Participants would have liked to see recommendations for improving eating habits based on the information entered. It seemed to be important to the participants to receive reminders for entering information about consumed food. Some participants mentioned that it would be helpful to be able to enter the exact amounts of food that was consumed to improve accuracy of the feedback. Participants liked the idea of creating their own goals, such as reaching a certain percentage of diversity across food groups, and they were interested in getting feedback on their target goals. They suggested receiving this feedback weekly or every few days. Almost all participants had difficulties understanding the meaning of the feedback around balanced diet and diet diversity and the meaning of the term *equilibrium*; they suggested to change the terminology or to provide an explanation.

Figure 2. Screenshot of the PEGASO Fit For Future eDiary app presented at the third iteration.



Implications for the PEGASO Fit for Future System Development

Table 5 summarizes the participants' design ideas for improving and changing the PEGASO F4F system components described in this paper and how the feedback was implemented for each technology. Quotes to support the findings are provided in Multimedia Appendix 5. Some of the design suggestions brought forward by the participants could not be implemented because of technical or practical reasons. For example, customizing the *challenges* function for each target behavior was deemed as too demanding in terms of time and resources for the development of the PEGASO F4F system. The option of exchanging points acquired by engaging with the PEGASO F4F system for rewards in the real world has been explored by the consortium but proved not to be viable at this stage of the intervention development.

For the development of the eDiary app, the SUS scores were used to track the evolution of the usability of the app and to ensure that the interface remained acceptable for participants. Changes made to the interface of the eDiary app in response to testing the app in iteration 2 (Table 5), yielded in improved SUS scores for the eDiary app from iteration 2 to iteration 3. This indicated that the eDiary app gained in usability and functionality. SUS scores for the Companion app were obtained in the third iteration only, and they indicated acceptable usability of the app.

SUS scores for the serious game indicated marginal usability for participants of 3 out of 4 test sites in the second iterations. SUS scores dropped in the third iteration in those 3 countries. Low SUS score for serious games are not uncommon in adolescents because they (1) implicitly compare serious games with commercial games [33] and (2) typically play action games on consoles or personal computers rather than mobile phones [34]. For this reason, the development of the serious game was primarily guided by focus group feedback.

Discussion

Principal Findings

This study aimed to describe the process of and findings from co-designing and prototyping components of a mobile phone-based technological system for promoting a healthy diet, physical activity, and reducing sedentary behavior in adolescents. Findings of this study suggested that for the PEGASO F4F technology to be motivating and engaging to adolescents, it should (1) allow personalization of the interface, (2) have age-appropriate and easy-to-understand language (of icons, labels, instructions, and notifications), (3) provide easily accessible tutorials on how to use the app or navigate through a game, (4) present a clear purpose and end goal, (5) have an appealing and self-explanatory reward systems, and (6) offer variation in gamified activities within apps and the serious game. Our findings also stress the importance of the ability of the mHealth technology to seek peer support and connect with peers for increased motivation and engagement with both mobile phone apps and the serious game through collaborative and competitive elements such as leader boards and joint activities. Incorporation of these user requirements led to the development of PEGASO F4F technology that combines 4 functionalities that are rarely united in currently existing mHealth technologies for healthy weight promotion in adolescents: (1) assisting the user in tracking and reviewing their behavior, (2) persuading the user to adopt intended behaviors through a serious game and gamification of apps, (3) advising the user on how to adopt intended behaviors, and (4) providing a platform for the user to interact with peers to encourage adoption of intended health behaviors. Therefore, PEGASO F4F advanced the current state

of the art and existing body of literature on mHealth for obesity prevention in young people.

Comparison With Prior Work

The current literature on the development of mHealth technology for prevention of obesity typically focused its efforts on a single app, game, or behavior (diet or physical activity), thereby ignoring the importance of promoting reduced sedentary behavior, and studies typically involved small numbers of participants in a single country [1,4,35,36]. PEGASO F4F is the first study that applied the principles of UCD of mHealth technology for adolescent healthy weight promotion at scale: co-design and feasibility testing of 3 different system components (2 apps and a serious game) for 3 health behaviors with 74 adolescents in 3 iterative cycles over the course of 12 months in 3 countries.

To date, studies reporting 3 or more cycles of iterative mHealth technology development for healthy weight promotion with adolescents are scarce; thus, the design of function and features of apps and games tend to rely on perceptions of the designers on users' needs and requirements [37]. The lack of iterative circles of technology design may result in a substantial level of uncertainty in usability, satisfaction, and acceptability of the new technology. With limited resources for research and concerns of producing ineffective interventions, intervention developers are urged to invest in thorough development of interventions involving the intended user populations before piloting the intervention and study design procedures in a controlled experimental study [38]. Taking the time and resources to incorporate adolescents' ideas, needs, and preference and present them with further developed technology twice, resulted in increased confidence of the PEGASO F4F consortium for producing an engaging mHealth technology for healthy weight promotion in a diverse group of adolescents. This was the case despite limited improvement of the SUS-obtained usability scores for the 3 apps reported in this paper. Although the eDiary app and serious game were integrated with the Companion app, other PEGASO F4F system components (eg, wearable technology and dashboard app) were not fully integrated during the third and final iteration of the co-design study. It is likely that this influenced the user experience and usability when engaging with the apps and serious game. In particular, the wearable technology was an important part for the functionality and user experience of the serious game linking real-world physical activity to the serious game (further explanations go beyond the scope of this paper). However, it is important to note that each co-design iteration involved different young people in most study sites and, as noted above, not all feedback and views could be incorporated. Therefore, the co-design process presented a series of compromises which was tested in the next phase. The next phase of the PEGASO F4F project tested the usability and acceptability of all fully integrated system component.

The approach taken for developing the PEGASO F4F system was in line with most research demonstrating end-user involvement (ie, as informants or co-designers). Therefore, involving adolescents in the development of the PEGASO F4F system as informants during 3 iterative focus group sessions in which they could verbalize their requirements of the system and ideas and feelings about the technology was considered a valid approach. However, it should be noted that results of a meta-analysis suggested that an active co-design role does not always increase serious game effectiveness as users often cannot relate the app or serious game to its learning objectives [28]. To combat this, the iterative stages included multiple ways for adolescents to (1) critically analyze their use and experience of the system, (2) provide feedback, and (3) contribute to the co-design process. By offering multiple methods to participants with differing capabilities and strengths, it increases the opportunity for each stage of technology development to capture valuable and accurate data from a population with wide intra- and intervariability [39]. Previous work on serious games has experienced similar variation in the output and value of end user involvement throughout the design process [40]. A concept introduced by Khaled and Vasalou [41] suggests a finer approach to participant selection in co-design, differentiating contributions in terms of domain expertise and procedural aspects to overcome some of the limitations that hinder current contributions to the design process; however, this approach has not been extensively tested. The authors consider that there is much more to learn about the application of participatory design and co-design methodology in the conceiving, design, development, and implementation of serious games.

Some design ideas could not be implemented for technical and practical reasons; however, this brings the research study in line with real-world development projects for which financial and time constraints are a common challenge. Lean development methods aim to overcome some of those challenges; however, in practice, research and development projects cannot always keep to these more rigid cycles because of their exploratory nature and work in innovative spaces.

Study Limitations

Focus groups and questionnaires were used for collecting design ideas and feedback on the usability and acceptability of the PEGASO F4F technology, and feedback guided the development of the apps and serious game. Although the use of focus groups and questionnaires reflects a solid approach to ascertain usability, additional methods need to be considered when refining the PEGASO F4F apps and serious game in future research.

First, the use of think aloud methods while adolescents are playing the game and using the apps will give more information about usability issues and the cause of these problems [42]. This will also minimize the chance of missing information because of recall bias, given that adolescents would play the game and verbalize feedback immediately instead of reporting experiences through questionnaires and interviews retrospectively.

Second, collection and interpretation of system usage data can provide valuable information in terms of measuring the level of engagement with mHealth technology [43]. A mixed method approach is recommended during mHealth technology development in which quantitative data (eg, back-end system data) reflect objective usage and qualitative data (eg, semistructured interview) provide more insight into reasons for playing or using, or more importantly, not playing the game or

using apps [43]. This would give designers and researchers a better and a more concrete understanding on which elements to adapt, and how to translate adolescents' feedback into engaging age-appropriate app and game design [5].

Third, to benefit from current efforts in gathering adolescents' feedback, it is recommended to get a designer involved during feedback sessions. As part of a multidisciplinary interview team, they would have a good understanding of game dynamics and ask appropriate follow-up questions that could be valuable for improved game design and possibly effectiveness. This would be in line with previous research results indicating that gathering feedback from end users should prioritize game dynamics (eg, level) above game aesthetics (eg, storyline and character) as this seems more strongly related to game effectiveness [40]. Involving game and app developers in the focus groups across the 3 countries and 4 intervention sites was considered for this study, but it was deemed not feasible because of constraints of time, proficiencies in local languages (ie, Catalan), and financial resources.

Another limitation of this study was that all PEGASO F4F prototypes were available in English language only. Although none of the adolescents approached to participate in this study (entire year groups) was excluded because of insufficient English language proficiency, participants in Spain and Italy stressed the importance of having apps available in local languages. The translation of the apps into Italian, Spanish, and Catalan was anticipated from the onset of the PEGASO F4F project and was done for testing the technology in the subsequent controlled pilot study. The lack of multilanguage prototypes might have impacted on participants' engagement with the technology. However, the design workshops and focus groups were conducted in the local language, allowing participants to express their creativity, ideas, and preferences for the development of the technology. Translation and back-translation of the apps and serious game content was not feasible during the technology development process because of the tight project timeline and constraints in the budget.

Study Implications and Lessons Learned

This study informed the PEGASO F4F system to be tested in a controlled multicenter pilot study. It provided 3 co-designed system components, the Companion app, eDiary. and serious game. In the PEGASO F4F controlled pilot study, adolescents across 4 European study sites tested the integrated system comprising the Companion app (including the Challenge function), eDiary app, Dashboard app, the serious game. and the wearable (smart garment and wrist-worn activity tracker). The individual system components of the pilot study were described in detail in the pilot study protocol [44].

The findings of this study might be of interest to other researchers and technology developers working in the field of mHealth intervention development for an adolescent population. Especially those developing mHealth interventions for promotion of physical activity and healthy eating in the context of obesity prevention or treatment could consider the following lessons learned:

- mHealth technology for adolescents should allow personalization of app interfaces, have instructions that describe the purpose of and navigation through the app, and offer rewards.
- The language used in the instructions and content should be age appropriate, and researchers should test the appropriateness of the language with the potential user. Adolescents particularly dislike repetitive tasks, and so mHealth technology should avoid unnecessary repetitions and offer variations in activities and content.
- For mHealth technology to be engaging, it should enable adolescents to connect with peers for support in adopting behavior change.

Conclusions

Incorporating adolescents' preferences, the PEGASO F4F technology combines the functions of a self-monitoring, entertainment, advisory, and social support tool. This study demonstrated that it is possible to employ an iterative co-design approach for development of a complex mHealth technology for adolescent healthy weight promotion in an international context. Participants' involvement over 3 iterative co-design circles informed the PEGASO F4F system to be tested in a controlled multicenter pilot study.

Acknowledgments

The authors thank all participants for their valuable feedback, creativity, and time. They are grateful for the support received by the participating schools. This project was funded by research grants from the European Commission under the Seventh Framework Programme (call identifier: FP7-ICT-2013-10; project number: 610727). AM was supported by the UK Medical Research Council (grant number MC_UU_12017/14) and the Scottish Government Chief Scientist Office (grant number SPHSU14). The funders had no involvement in review and approval of the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Companion app and eDiary app co-design worksheet.
[PDF File (Adobe PDF File), 15564 KB - mhealth_v8i3e14118_app1.pdf]

Multimedia Appendix 2
Focus group discussion guides.
[DOCX File , 19 KB - mhealth_v8i3e14118_app2.docx]

Multimedia Appendix 3
The brief use of app questionnaire.
[DOCX File , 14 KB - mhealth_v8i3e14118_app3.docx]

Multimedia Appendix 4
Serious game demo video.
[MP4 File (MP4 Video), 14897 KB - mhealth_v8i3e14118_app4.mp4]

Multimedia Appendix 5
Participants' quotes.
[DOC File , 49 KB - mhealth_v8i3e14118_app5.doc]

References

1. Turner T, Spruijt-Metz D, Wen CK, Hingle MD. Prevention and treatment of pediatric obesity using mobile and wireless technologies: a systematic review. Pediatr Obes 2015 Dec;10(6):403-409 [FREE Full text] [doi: 10.1111/ijpo.12002] [Medline: 25641770]
2. Hekler EB, Michie S, Pavel M, Rivera DE, Collins LM, Jimison HB, et al. Advancing models and theories for digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):825-832. [doi: 10.1016/j.amepre.2016.06.013] [Medline: 27745682]
3. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. Transl Behav Med 2016 Jun;6(2):317-328 [FREE Full text] [doi: 10.1007/s13142-016-0395-7] [Medline: 27357001]
4. Dute DJ, Bemelmans WJ, Breda J. Using mobile apps to promote a healthy lifestyle among adolescents and students: a review of the theoretical basis and lessons learned. JMIR Mhealth Uhealth 2016 May 5;4(2):e39 [FREE Full text] [doi: 10.2196/mhealth.3559] [Medline: 27150850]
5. Baranowski T, Blumberg F, Buday R, DeSmet A, Fiellin LE, Green CS, et al. Games for health for children-current status and needed research. Games Health J 2016 Feb;5(1):1-12 [FREE Full text] [doi: 10.1089/g4h.2015.0026] [Medline: 26262772]
6. Berger V, Schrader U. Fostering sustainable nutrition behavior through gamification. Sustainability 2016;8(1):67 [FREE Full text] [doi: 10.3390/su8010067]
7. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
8. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
9. Goodyear VA, Armour KM. Young people’s perspectives on and experiences of health-related social media, apps, and wearable health devices. Soc Sci 2018;7(8):137. [doi: 10.3390/socsci7080137]
10. Jensen CD, Duncombe KM, Lott MA, Hunsaker SL, Duraccio KM, Woolford SJ. An evaluation of a smartphone-assisted behavioral weight control intervention for adolescents: pilot study. JMIR Mhealth Uhealth 2016 Aug 23;4(3):e102 [FREE Full text] [doi: 10.2196/mhealth.6034] [Medline: 27554704]
11. Lyzwinski LN, Caffery L, Bambling M, Edirippulige S. The Mindfulness app trial for weight, weight-related behaviors, and stress in university students: randomized controlled trial. JMIR Mhealth Uhealth 2019 Apr 10;7(4):e12210 [FREE Full text] [doi: 10.2196/12210] [Medline: 30969174]
12. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. Co-design 2008;4(1):5-18. [doi: 10.1080/15710880701875068]
13. Steen M, Manschot M, de Koning N. Benefits of co-design in service design projects. Int J Des 2011;5(2):53-60 [FREE Full text]
14. Craven M, Lang A, Martin J. Developing mHealth Apps With Researchers: Multi-Stakeholder Design Considerations. In: Proceedings of the International Conference of Design, User Experience, and Usability. Switzerland: DUXU 2014. Lecture Notes in Computer Science, vol 8519. Springer, Cham; 2014 Presented at: DUXU'14; June 22-27, 2014; Heraklion, Crete, Greece. [doi: 10.1007/978-3-319-07635-5]

15. Eyles H, Jull A, Dobson R, Firestone R, Whittaker R, Te Morenga L, et al. Co-design of mHealth delivered interventions: a systematic review to assess key methods and processes. Curr Nutr Rep 2016;5(3):160-167. [doi: 10.1007/s13668-016-0165-7]
16. Whitehouse SR, Lam PY, Balka E, McLellan S, Deevska M, Penn D, et al. Co-creation with TickiT: designing and evaluating a clinical eHealth platform for youth. JMIR Res Protoc 2013 Oct 18;2(2):e42 [FREE Full text] [doi: 10.2196/resprot.2865] [Medline: 24140595]
17. Carrino S, Caon M, Angelini L, Mugellini E, Khaled OA, Orte S, et al. PEGASO: A Personalised and Motivational ICT System to Empower Adolescents Towards Healthy Lifestyles. Stud Health Technol Inform 2014;207:350-359. [Medline: 25488241]
18. Carrion C, Moliner LA, Castell C, Puigdomènech E, Gómez FS, Domingo L, et al. [Use of the Smartphone to Promote Healthy Habits among Teen-agers, Spain]. Rev Esp Salud Publica 2016 Nov 3;90:e1-11 [FREE Full text] [Medline: 27808084]
19. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016 Dec 16;18(12):e317 [FREE Full text] [doi: 10.2196/jmir.5927] [Medline: 27986647]
20. Lang AR, Craven MP, Atkinson S, Simons L, Cobb S, Mazzola M. Human factors multi-technique approach to teenage engagement in digital technologies health research. In: Little L, Fitton D, Bell BT, Toth N, editors. Perspectives on HCI Research with Teenagers. New York City: Springer International Publishing; 2016:61-101.
21. Velickovski F, Orte S, Sola M, Tabozzi SA, Lafortuna CL. Detection and assessment of behaviours associated with the risk of obesity in adolescents. In: Giokas K, Bokor L, Hopfgartner F, editors. eHealth 360°. Switzerland: Springer; 2016:253-258.
22. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
23. Seligman ME, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. Am Psychol 2005;60(5):410-421. [doi: 10.1037/0003-066X.60.5.410] [Medline: 16045394]
24. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
25. Thaler RH, Sunstein CR. Nudge: Improving Decisions About Health, Wealth, and Happiness. Westminster, London: Penguin Books; 2009.
26. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification. Using Game-design Elements in Non-gaming Contexts. In: Proceedings of the Extended Abstracts on Human Factors in Computing Systems. New York: Association for Computing Machinery; 2011 Presented at: CHI'11; May 7 - 12, 2011; Vancouver, Canada p. 2425-2428. [doi: 10.1145/1979742.1979575]
27. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. J Med Internet Res 2016 Jan 8;18(1):e6 [FREE Full text] [doi: 10.2196/jmir.4790] [Medline: 26747176]
28. Boyle EA, Hainey T, Connolly TM, Gray G, Earp J, Ott M, et al. An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Comput Educ 2016;94:178-192. [doi: 10.1016/j.compedu.2015.11.003]
29. Mesas AE, Muñoz-Pareja M, López-García E, Rodríguez-Artalejo F. Selected eating behaviours and excess body weight: a systematic review. Obes Rev 2012 Feb;13(2):106-135. [doi: 10.1111/j.1467-789X.2011.00936.x] [Medline: 21955734]
30. Brooke J. SUS - A quick and dirty usability scale. Usability Eval industry 1996;189(194):4-7 [FREE Full text] [doi: 10.1201/9781498710411]
31. Albert W, Tullis T. Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics. Second Edition. Burlington, Massachusetts: Morgan Kaufmann; 2013.
32. Nielsen J. Enhancing the Explanatory Power of Usability Heuristics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery; 1994 Presented at: CHI'94; April 24 - 28, 1994; Massachusetts, Boston, USA p. 152-158. [doi: 10.1145/191666.191729]
33. DeVary S. Educational gaming: interactive edutainment. Dist Learn 2008;5(3):35 [FREE Full text]
34. Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM. A systematic literature review of empirical evidence on computer games and serious games. Comput Educ 2012;59(2):661-686. [doi: 10.1016/j.compedu.2012.03.004]
35. Baretta D, Perski O, Steca P. Exploring users' experiences of the uptake and adoption of physical activity apps: longitudinal qualitative study. JMIR Mhealth Uhealth 2019 Feb 8;7(2):e11636 [FREE Full text] [doi: 10.2196/11636] [Medline: 30735143]
36. Rohde A, Duensing A, Dawczynski C, Godemann J, Lorkowski S, Brombach C. An app to improve eating habits of adolescents and young adults (Challenge to Go): systematic development of a theory-based and target group-adapted mobile app intervention. JMIR Mhealth Uhealth 2019 Aug 12;7(8):e11575 [FREE Full text] [doi: 10.2196/11575] [Medline: 30903746]
37. Hermawati S, Lawson G. Managing obesity through mobile phone applications: a state-of-the-art review from a user-centred design perspective. Pers Ubiquit Comput 2014;18(8):2003-2023. [doi: 10.1007/s00779-014-0757-4]

38. Hoddinott P. A new era for intervention development studies. Pilot Feasibility Stud 2015;1:36 [FREE Full text] [doi: 10.1186/s40814-015-0032-0] [Medline: 27965814]
39. Lang AR, Martin JL, Sharples S, Crowe JA. The effect of design on the usability and real world effectiveness of medical devices: a case study with adolescent users. Appl Ergon 2013 Sep;44(5):799-810. [doi: 10.1016/j.apergo.2013.02.001] [Medline: 23453773]
40. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, de Bourdeaudhuij I. Is participatory design associated with the effectiveness of serious digital games for healthy lifestyle promotion? A meta-analysis. J Med Internet Res 2016 Apr 29;18(4):e94 [FREE Full text] [doi: 10.2196/jmir.4444] [Medline: 27129447]
41. Khaled R, Vasalou A. Bridging serious games and participatory design. Int J Child-Comput Interact 2014;2(2):93-100. [doi: 10.1016/j.ijcci.2014.03.001]
42. Donker A, Markopoulos P. A comparison of think-aloud, questionnaires and interviews for testing usability with children. In: Faulkner X, Finlay J, Détienne F, editors. People and Computers XVI-Memorable Yet Invisible. London: Springer; 2002:305-316.
43. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
44. Puigdomenech E, Martin A, Lang A, Adorni F, Gomez SF, McKinstry B, PEGASO Fit For Future Consortium. Promoting healthy teenage behaviour across three European countries through the use of a novel smartphone technology platform, PEGASO fit for future: study protocol of a quasi-experimental, controlled, multi-Centre trial. BMC Med Inform Decis Mak 2019 Dec 17;19(1):278 [FREE Full text] [doi: 10.1186/s12911-019-0958-x] [Medline: 31847919]

Abbreviations

F4F: Fit for Future
HCI: human-computer interaction
mHealth: mobile health
SUS: System Usability Scale
UCD: user-centered design

Edited by G Eysenbach; submitted 25.03.19; peer-reviewed by K Stasiak, S Smith, N Vu; comments to author 29.05.19; revised version received 20.08.19; accepted 16.12.19; published 02.03.20.

Please cite as:
Martin A, Caon M, Adorni F, Andreoni G, Ascolese A, Atkinson S, Bul K, Carrion C, Castell C, Ciociola V, Condon L, Espallargues M, Hanley J, Jesuthasan N, Lafortuna CL, Lang A, Prinelli F, Puidomenech Puig E, Tabozzi SA, McKinstry B
A Mobile Phone Intervention to Improve Obesity-Related Health Behaviors of Adolescents Across Europe: Iterative Co-Design and Feasibility Study
JMIR Mhealth Uhealth 2020;8(3):e14118
URL: https://mhealth.jmir.org/2020/3/e14118
doi:10.2196/14118
PMID:32130179

Original Paper

Supporting the Medication Adherence of Older Mexican Adults Through External Cues Provided With Ambient Displays: Feasibility Randomized Controlled Trial

Ernesto Zárate-Bravo[1*], MSc; Juan-Pablo García-Vázquez[1], PhD; Engracia Torres-Cervantes[2], BSCN; Gisela Ponce[2], PhD; Ángel G Andrade[1], PhD; Maribel Valenzuela-Beltrán[1], BSc; Marcela D Rodríguez[1*], PhD[‡]

[1]Facultad de Ingeniería, Universidad Autónoma de Baja California, Mexicali, Mexico

[2]Facultad de Enfermería, Universidad Autónoma de Baja California, Mexicali, Mexico

[‡]Written on behalf of Academic Groups on Intelligent Systems and Community Health

[*]these authors contributed equally

Corresponding Author:
Marcela D Rodríguez, PhD
Facultad de Ingeniería
Universidad Autónoma de Baja California
Ave Alvaro Obregón, S/N
Col Nueva
Mexicali, 21100
Mexico
Phone: 52 6865664270 ext 1301
Email: marcerod@uabc.edu.mx

Abstract

Background: Problems with prospective memory, which refers to the ability to remember future intentions, cause deficits in basic and instrumental activities of daily living, such as taking medications. Older adults show minimal deficits when they rely on mostly preserved and relatively automatic associative retrieval processes. On the basis of this, we propose to provide external cues to support the automatic retrieval of an intended action, that is, to take medicines. To reach this end, we developed the Medication Ambient Display (MAD), a system that unobtrusively presents relevant information (unless it requires the users' attention) and uses different abstract modalities to provide external cues that enable older adults to easily take their medications on time and be aware of their medication adherence.

Objective: This study aimed to assess the adoption and effect of external cues provided through ambient displays on medication adherence in older adults.

Methods: A total of 16 older adults, who took at least three medications and had mild cognitive impairment, participated in the study. We conducted a 12-week feasibility study in which we used a mixed methods approach to collect qualitative and quantitative evidence. The study included baseline, intervention, and postintervention phases. Half of the participants were randomly allocated to the treatment group (n=8), and the other half was assigned to the control group (n=8). During the study phases, research assistants measured medication adherence weekly through the pill counting technique.

Results: The treatment group improved their adherence behavior from 80.9% at baseline to 95.97% using the MAD in the intervention phase. This decreased to 76.71% in the postintervention phase when the MAD was no longer being used. Using a one-way repeated measures analysis of variance and a post hoc analysis using the Tukey honestly significant difference test, we identified a significant statistical difference between the preintervention and intervention phases (P=.02) and between the intervention and postintervention phases (P=.002). In addition, the medication adherence rate of the treatment group (95.97%) was greater than that of the control group (88.18%) during the intervention phase. Our qualitative results showed that the most useful cues were the auditory reminders, followed by the stylized representations of medication adherence. We also found that the MAD's external cues not only improved older adults' medication adherence but also mediated family caregivers' involvement.

Conclusions: The findings of this study demonstrate that using ambient modalities for implementing external cues is useful for drawing the attention of older adults to remind them to take medications and to provide immediate awareness on adherence behavior.

Trial Registration: ClinicalTrials.gov NCT04289246; https://tinyurl.com/ufjcz97

(JMIR Mhealth Uhealth 2020;8(3):e14680) doi:10.2196/14680

KEYWORDS

health information systems; family caregiver; aged; medication adherence

Introduction

Background

One of the most common reasons for medication nonadherence among older adults is forgetfulness [1,2]. Problems with prospective memory, that is, the ability to remember future intentions, cause deficits in basic and instrumental activities of daily living, such as taking medications [3]. Consequently, the responsibility of managing medications of older adults and people living with dementia often falls on family caregivers [4-6]. Therefore, older adults need medication management technologies to help them not only with forgetfulness or cognitive impairment but also with other reasons that contribute to the lack of adherence to medications, such as polypharmacy (ie, the self-administration of multiple drugs) [2]. Our work is based on the facts that older adults show "substantial deficits when they depend on working memory and executive resources for prospective recall, but minimal deficits when they rely on associative recovery processes" [1]. Research has also shown that the use of external cues supports the automatic recovery of a planned action [1-10]. Therefore, we propose to support external cues through a tablet-based ambient display designed to increase the retrieval process of the planned action (ie, taking medications) and to provide awareness of adherence behavior [11,12]. To reach this end, we designed the Medication Ambient Display (MAD) [12,13].

An ambient display unobtrusively presents relevant information unless it requires the users' attention [14]. In addition, users can easily monitor the display to obtain the desired information because it uses abstract modalities to represent information, such as pictures, sounds, and movement [14]. Thus, we used different abstract modalities to provide external cues that enable older adults to easily obtain relevant information to take their medications on time and be aware of their medication adherence.

Previous Works and Study Rationale

In the last decade, different technological-based interventions for supporting the medication adherence of older adults have been studied. Owing to the rapid penetration of mobile phones, one of the approaches that has been widely studied is text message (SMS) reminders [15]. A systematic review identified that 18 of the 29 selected studies reported SMS-based interventions that improved older adults' medication adherence [16]. A review of commercially available medication management apps for mobile phones reported that most of them focus on providing reminders [17]. A qualitative study identified that an interaction supported through linear navigation and multimodal reminder methods should be considered to increase the ease of use of mobile medication apps available on the internet; thus, these could be adopted by older adults [18]. Similarly, an adoption assessment of a commercial telehealth medication-dispensing device found that this is an acceptable tool for older adults to manage medications in collaboration with nurses [19].

Some research has been conducted to explore new computing approaches, such as ambient computing technologies, mobile games, and conversational agents specifically designed for older adults [20-23]. Several works have focused on supporting persuasive strategies to motivate older adults to follow their medication regimens. For instance, MoviPill is a mobile phone app that gamifies medication activity by awarding seniors who take their daily medication doses at the prescribed time and by promoting social competition [21]. dwellSense is a peripheral display that provides real-time and explicit feedback about medication-taking behaviors (eg, what medications were taken and whether they were taken on time, late, or not taken at all) [20]. This medication feedback encouraged seniors to reflect on their medication errors and then improve their medication self-efficacy [20]. Recently, the use of conversational agents accessible from mobile devices was explored to support educational strategies. The Conversational Medication Assistant for Heart Failure, named CARMIE, was developed to provide older adults with advice and information about their medication regimens, such as medication interactions, adverse reactions, and indications about how to take medications [23]. Similarly, the Electronic Medication Management Assistant is a chatbot developed to coach patients for managing medications prescribed by different health care providers, such as detecting double prescriptions, medication interactions, and contraindications [22]. In contrast to the previously mentioned works, our technological approach uses ambient modalities to provide external cues that aim to increase the retrieval process of the planned action and to provide daily and immediate awareness about how medication regimens are followed during the day. We assessed the effect of our approach by using objective medication adherence measures; moreover, we obtained qualitative findings that help us understand the adoption of the MAD. For this end, we provided seniors with the MAD to support the medication treatments prescribed by their physicians and the timetables that the seniors themselves proposed to follow.

Objectives

Our study aimed to address the following research questions (RQ):

> *RQ1: What is the effect of the external cues provided by the MAD on older adults' medication adherence?*
>
> *RQ2: How do the MAD design features promote its adoption?*

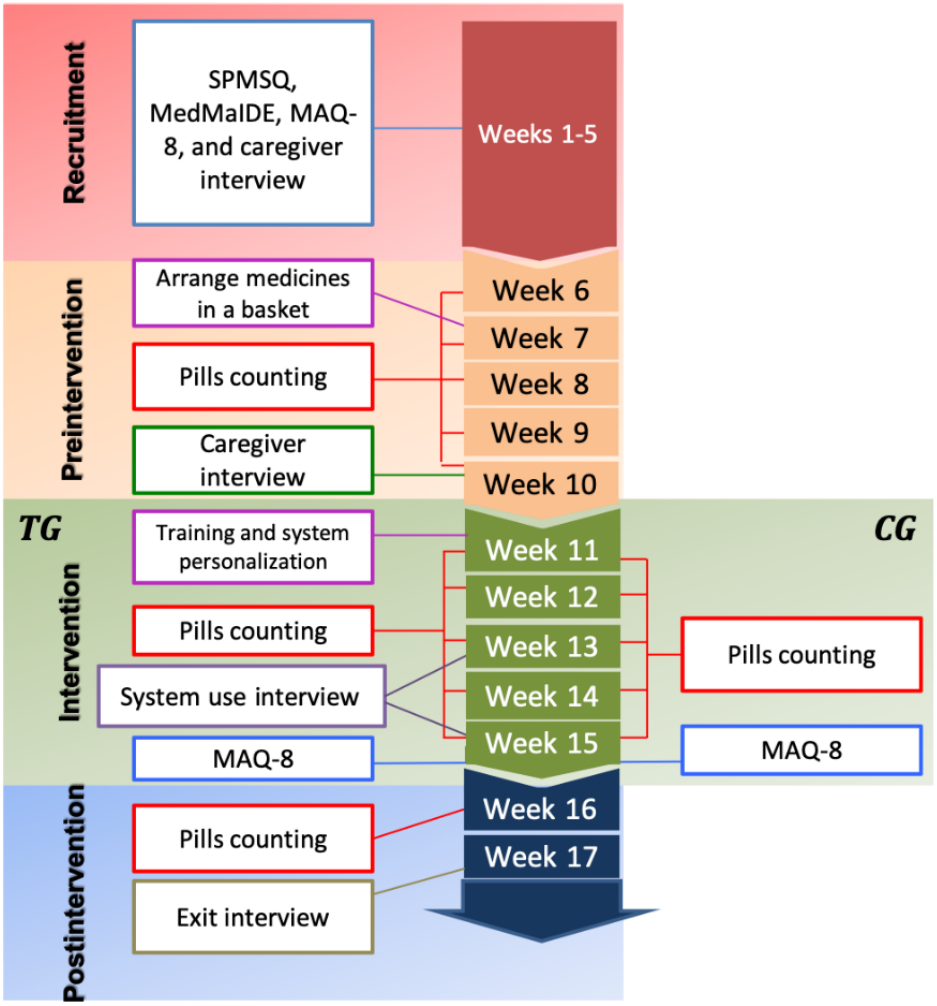
We used a mixed methods approach to obtain quantitative evidence regarding how several variables associated with adherence to medication improved and qualitative evidence

regarding the adoption of the system by the participants. To describe our study, we used the suggestions by the Consolidated Standards of Reporting Trials to report pilot investigations [24] and electronic health interventions [25].

Methods

The evaluation of the MAD was designed as a small trial study. This section presents the study timeline, the activities conducted, and the instruments used to collect data (see Figure 1).

Figure 1. Study activities and instruments administered to participants during each study phase. CG: control group; MAQ-8: 8-item Medication Adherence Questionnaire; MedMaIDE: Medication Management Instrument for Deficiencies in the Elderly; SPMSQ: Short Portable Mental Status Questionnaire; TG: treatment group.



Participants Recruitment

Eligibility Criteria

To be eligible, older adults had to meet the following criteria: be older than 60 years, take at least three medications prescribed by a physician (ie, polypharmacy), have mild cognitive impairment, report medication-forgetting events, and live with a relative who could provide us with information on the assistance required by the study participant to take their medications. The exclusion criteria were as follows: being unable to self-administer medications due to a functionality problem or severe cognitive impairment, and not taking pill-based medications (it may be difficult to assess adherence otherwise). To participate in the study, it was not a requirement that older adults have experience in the use of the internet or mobile devices.

Recruitment Procedure

The study was conducted in Mexicali, Mexico. Ten students from the Faculty of Nursing at the Universidad Autónoma de Baja California participated as research assistants. These students were enrolled in a social service program at the Community Center of the University (known as the UNICOM), which aims to provide seniors with occupational therapy and provide some health care assistance. The UNICOM is strategically located in a neighborhood where aging inhabitants predominate. For recruiting participants, research assistants contacted older adults in the vicinity of the UNICOM and administered a set of instruments as summarized in Table 1. First, the Spanish version of the 10-item Short Portable Mental Status Questionnaire was administered to detect the presence of mild cognitive impairment in the contacted seniors [26]. Then, the 8-item Medication Adherence Questionnaire (MAQ-8) scale was used to identify nonadherent seniors through 8 questions [27]. We selected this instrument because it is the

quickest scale to administer, the simplest to score, and has been validated in many populations with different diseases and in persons with low literacy [28]. Finally, the Medication Management Instrument for Deficiencies in the Elderly (MedMaIDE) instrument was used to assess if the participants had deficiencies in managing medications and whether they met the polypharmacy criteria [29]. The MedMaIDE was designed to be administered in the home setting by nonclinical experts. It consists of 20 items to find out what patients know about their medications, whether they know how to take it, and how to get it from a doctor or pharmacy. During the MedMaIDE administration, relatives were permitted to participate to complement the seniors' responses; this is considered appropriate to obtain a more reliable medication assessment [30]. In addition, relatives of older adults were interviewed to identify their role in helping older adults follow their medication routine. Older adults who met the eligibility criteria and expressed their interest to participate were enrolled in the study. The recruitment procedure lasted approximately 5 weeks.

Table 1. Instruments used to assess the eligibility criteria.

Eligibility criteria	Instrument	Score to be eligible
Mild cognitive impairment	Short Portable Mental State Questionnaire [26]	3 or 4 points
Medication deficiency	Medication Management Instrument for Deficiencies in the Elderly [29]	<13 points
Adherence for medicating	8-item Medication Adherence Questionnaire, also known as Morisky scale [27]	1-2 points=low and 3-8 points=medium
Caregiver involvement	Semistructured interview to find out how caregivers assisted older adults	—[a]

[a]Not applicable.

Organizational Setting

Once the participants were recruited, we realized that they were primarily of low socioeconomic status and affiliated with the Mexican Institute of Social Security (IMSS), the largest medical institution in Mexico. Periodically (monthly or bimonthly), they attended an IMSS clinic for follow-up consultation and to retrieve an updated prescription to get their medications from the clinic's pharmacy. The lack of adequate health care and pharmaceutical policies to rationally manage medications and monitor the treatment of patients increases the vulnerability of Mexican seniors to medication errors [31], a situation that is also faced in other countries [32,33].

Preintervention

Baseline data were collected during weeks 6 to 10 on medication adherence by using the pill counting technique. We noticed that participants accumulated containers with the same medications. Under those circumstances, we provided seniors with a basket to arrange the medications that should be taken each week (see Figure 2), which facilitated data collection for measuring the $Dosage_{pill}$ adherence outcome (see Table 2).

Figure 2. A research assistant arranging medications in a basket.

Table 2. Outcome variables and methods used to collect data to address research question 1.

Variable	Description	Collection method
$Dosage_{pill}$ [a]	The number of pills taken by participants in a period divided by the number of pills expected to be taken for that period[a]	Pill counting
$Dosage_{MAD}$ [b]	The number of medication episodes reported as taken by participants in a period divided by the number of episodes expected to be recorded for that period	MAD's[c] log
Timely[d]	Indicates whether the medication was taken 30 min before or after the time expected to take the medication. This is the number of medication episodes registered in the time window during a period divided by the number of episodes registered as taken for that period	MAD's log
Self-reported medication adherence[e]	A score estimated based on reported nonadherent behaviors; for example, drug omissions, medication forgetting, carelessness, or stopping a medication when feeling worse [28].	8-item Medication Adherence Questionnaire

[a]It measures whether the medication is not being taken as prescribed, which may affect the clinical outcome. We estimated it for both groups (treatment group and control group).

[b]It enabled us to understand how much older adults used the MAD reminders. It was estimated for the treatment group.

[c]MAD: Medication Ambient Display.

[d]It measures whether doses were taken during the prescribed interval. It was estimated for the treatment group.

[e]It is an assessment instrument to identify individuals' perception about their medication adherence. The 8-item Medication Adherence Questionnaire was administered during the recruitment phase and at the end of the intervention phase.

Intervention

Design and Implementation of the Medication Ambient Display

For medical information systems to be more specific to the needs of users, an iterative approach must be followed, which consists of different usability studies [34]. In this sense, our previous work included usability studies that helped us inform the design of the MAD. They included (1) a usability inspection in which experts (usability engineers and geriatricians) determined how the MAD conformed to usability design principles for ambient displays [11] and (2) a field evaluation to identify technical and usability problems that we addressed without altering the conceptual design of the MAD [12,13]. In this paper, we present the trial study that we conducted to evaluate the effect of the MAD on measures of adherence to medication and its possible adoption.

We implemented the MAD for Android tablets to be placed as portrait frames in the older adults' homes and to provide the following cues:

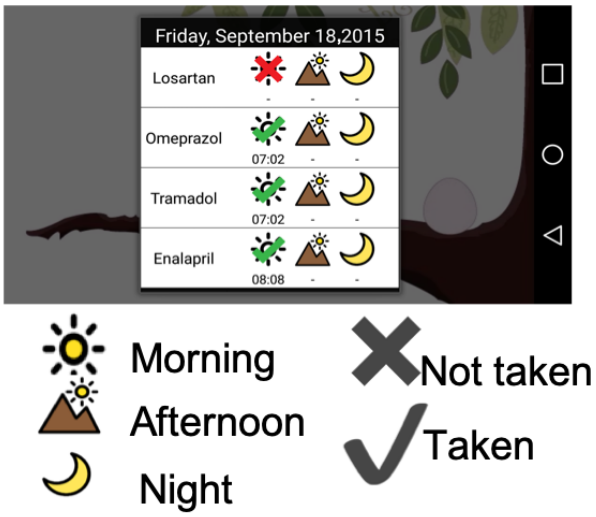
- Abstract and stylized representations of their medication adherence
- Auditory and visual reminders to call older adults' attention
- Events that may enhance older adults' awareness about whether the medication was taken

Abstract and Stylized Representations of Medication Adherence

The MAD shows a virtual birdcage, which has the aim of raising elders' consciousness about how they have to take responsibility for caring for their health, in a way similar to willingly caring for a pet. As presented in Figure 3, the abstract representation is an animation of a parakeet that symbolizes daily medication compliance. Each day, a newborn pet grows to represent medication compliance. In addition, by touching any point on the virtual cage of the parakeet, the MAD presents detailed information on the individual's daily medication compliance by using the notation presented at the bottom of Figure 4. In this figure, the MAD shows that an older adult has to take 4 medicines, and each of them should be taken 3 times during the day: morning, afternoon, and night. Thus, it presents that a morning medicine (Losartan) was not taken and that the afternoon and night doses are still pending. Optionally, participants can consult their medication adherence from any other date.

Figure 3. Abstract representations of medication adherence based on parakeet growth.

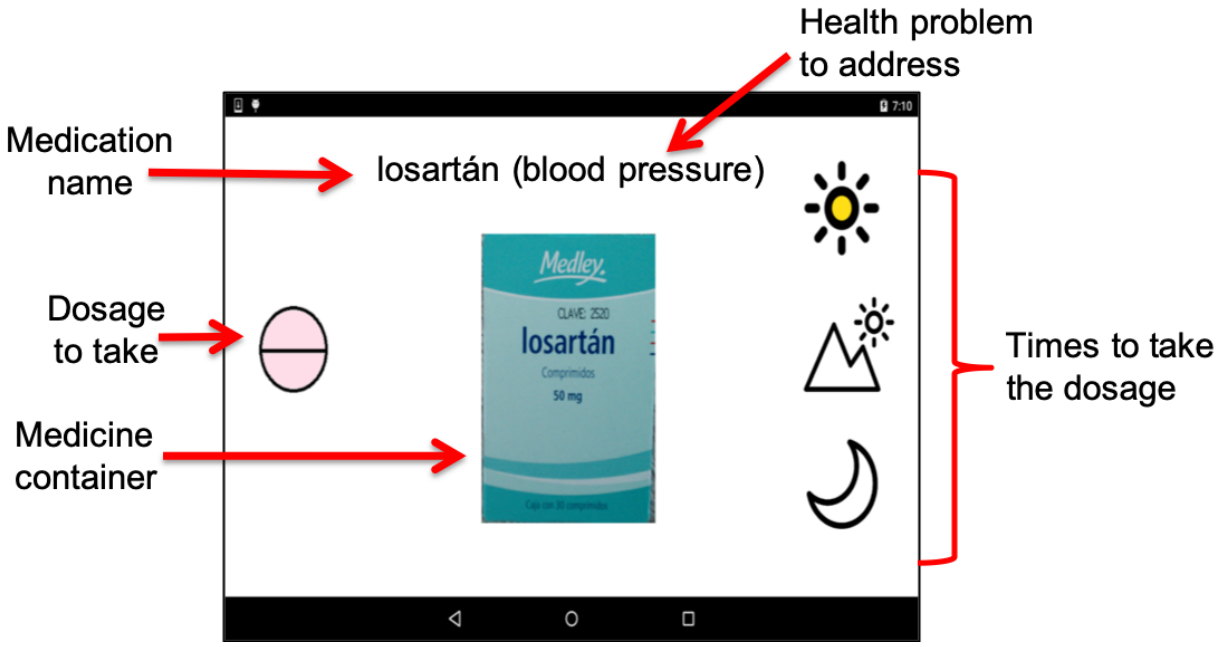
Figure 4. Detailed information about the medication adherence corresponding to the current day and notation used to represent if medicines were taken on time.



Auditory and Visual Reminders to Call Older Adults' Attention

The parakeet provides auditory reminders (ie, parakeet whistle) and pictograms that inform how to take medications. For instance, Figure 5 denotes the morning doses (ie, 1 pill of Losartan for controlling blood pressure). On the right side, the pictogram depicts through representative icons that this medication has to be taken 3 times a day: morning, afternoon, and night. Moreover, the morning icon is colored to denote that the MAD reminder is for the first doses of the day.

Figure 5. Medication Ambient Display reminding to medicate.



Events to Enhance Older Adults' Awareness of Whether the Medication Was Taken

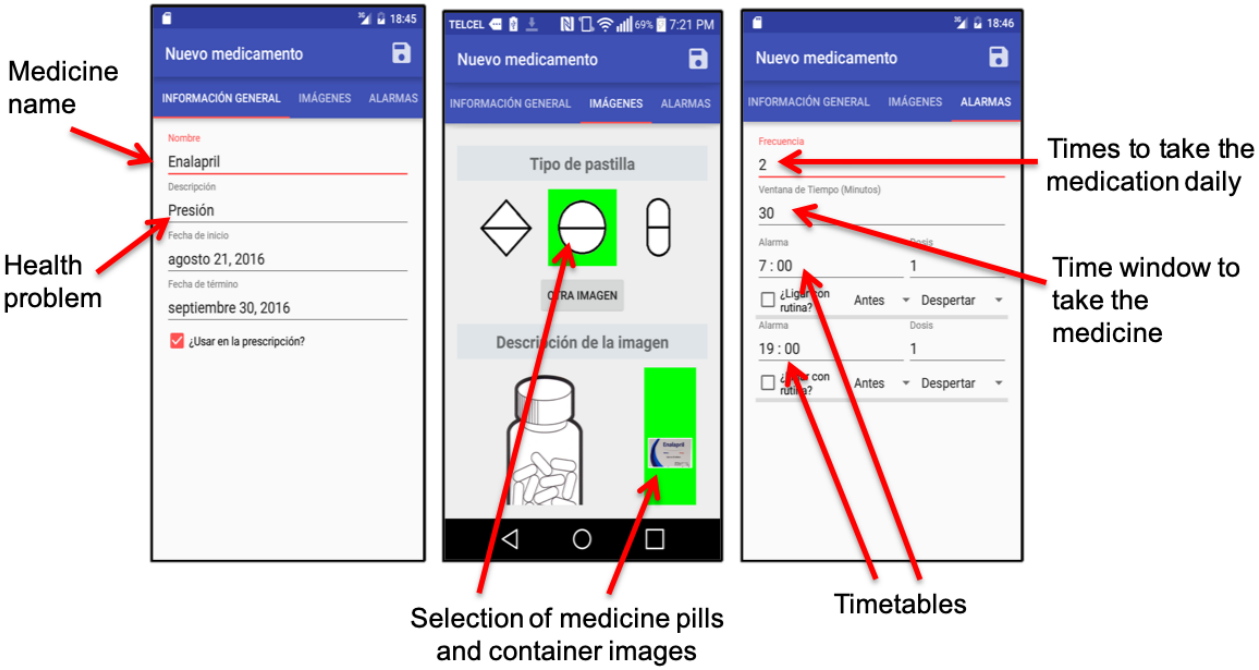
These cues refer to actions performed by older adults to make taking medications more memorable [12]. As depicted in Figure 6, after an individual takes their medication, they move the tablet closer to the pill container to indicate that the medication was taken. We implemented this functionality through Near Field Communication (NFC) technology. Afterward, the parakeet acknowledges that the medication was registered as taken.

We also implemented an administration component (see Figure 7), which we used to tailor MAD to participants' medication prescriptions and the timetables that they reported to follow. As evaluating the ease of adapting and configuring MAD was out of the scope of this study, participants did not use the MAD administrator. Finally, MAD also includes a component to generate a log with the medication episodes registered by the participants through the NFC technology. A medication episode

comprises the medication name and the corresponding timestamp. This information could be consulted during our visits to the seniors via this administrator component.

Figure 6. Registering the medication: After a senior medicates, she/he should move the corresponding medication container closer to the tablet in order for the attached NFC tag can be recognized by the tablet NFC reader (left); then, MAD acknowledges that the medicine was registered as taken on time (right).

Figure 7. User interfaces of the administration component of Medication Ambient Display, which shows how it enabled the information registration of each medication.



Intervention Procedure

We conducted a session with the research assistants, who made a random and blind allocation of the participants to the treatment group (TG) and the control group (CG). On the first day of this phase, research assistants visited older adults in the TG to introduce the MAD in the presence of caregivers by using the *spaced retrieval* approach, that is, teach, ask, wait, ask again, wait, and ask again [35]. This approach has been used to support the encoding, retention, and retrieval processes involved in interventions designed to assist medication taking [35]. Thus, we explained to participants how to carry out medication taking using the system. Afterward, we asked them to recall the system features we had just presented and assisted them as necessary. To do this, we activated each of the system's functionalities and asked participants to use them (eg, interpreting a reminder, registering a medication, and consulting and interpreting their medication compliance). Then, we waited for 1 min and asked them to recall the system's functionalities again. We asked them again 15 min later and repeated it once again. After the completion of the training session, which lasted 40 to 60 min approximately, the MAD was personalized according to the participant's prescriptions and through discussions with the participant on an appropriate schedule for presenting the reminders. Research assistants used the MAD administrator to enter the medication names, health problems to address, timetables, and the frequency at which medications should be taken. Then, they attached and configured NFC tags to each of the pill containers, which included selecting the images that

best represent the pills and their containers. Afterward, the MAD was placed in the area of participant's home where they usually reported taking medications, mostly the kitchen, living room, and bedroom. The intervention phase lasted 5 weeks (see Figure 1), during which research assistants visited participants to collect data on medication adherence (from both the TG and the CG) and system usage (from the TG).

Postintervention

After the intervention was completed, we removed the MAD from the participants' homes. Research assistants then carried out weekly visits (weeks 16-17) to older adults from the TG to collect data that enabled us to understand how the withdrawal of the MAD affected their medication routine and adherence.

Outcome Measures and Data Acquisition

Medication compliance (known as adherence as well) refers to "the act of conforming to the recommendations made by the provider concerning timing, dosage, and frequency of medication-taking" [27]. On the basis of this definition, we identified a set of variables as relevant for analyzing the effect of the external cues provided by the MAD on the participants' medication adherence (see Table 2). Thus, these variables were used to address RQ1.

During the intervention phase, we also collected qualitative evidence about the system's adoption, which enabled us to address RQ2. We interviewed the older adults in the TG regarding the system's functionalities that they perceived as most useful, less useful, and the difficulties faced while using it. At the end of the postintervention stage, we interviewed participants to obtain their perceptions of how withdrawal from the MAD impacted their medication adherence. In addition, those caregivers who were at home during our visits were interviewed to obtain information on their involvement in the seniors' medication activities. Our questions centered on the specific activities associated with the older adult's medication regimen that caregivers were involved in and how they knew if the older adult took his or her pills in a given week. These semistructured interviews were administered by the first 3 authors of this paper.

Data Analysis

We used Student *t* tests and chi-square tests to measure the statistical difference in age, the number of prescribed medicines, and self-reported medication adherence between the TG and the CG. A one-way repeated measure analysis of variance (ANOVA), dependent *t* tests, and independent *t* tests were used to find significant differences in medication adherence between the study phases and between the TG and the CG. The McNemar test was used to verify differences within the TG between the self-reported medication adherence in the recruitment and intervention phases. To determine whether any of the differences between the means estimated are statistically significant, we compared the *P* value with a significance level set to .05 [35].

For the qualitative analysis, we transcribed the collected data from their original Spanish version, that is, audio, handwritten notes, and photographs taken during the interviews. Individual quotes were translated into English for use in this paper. We followed the thematic analysis approach, which consists of generating initial codes from the data, searching for potential themes, contrasting the identified themes with the data, and iteratively refining them [36].

Ethics Statement

The ethics review board of Faculty of Nursing approved the study protocol once we proposed how to address their suggestions on how to handle the withdrawal of the technology at the study end. We agreed to provide the participants of the TG with an adequate financial incentive that would allow them (if desired) to obtain a PC tablet similar to the one used during the study. Every week, participants received an economic incentive, approximately US $7 if they were in the CG and US $14 if they were in the TG. We obtained informed written consent from all individual participants.

Results

Baseline Characteristics of Study Participants

The research assistants contacted approximately 100 older adults to participate in the study (see Figure 8). They identified 42 potential participants; 20 of them met the eligibility criteria and were enrolled in the study. However, only 16 completed the study. The analysis of baseline data presented in Table 3 indicated that the TG and the CG had no significant differences in age (*P*=.21), number of medicines taken (*P*=.33), self-reported medication adherence (*P*=.59), education in years (*P*=.35), gender (*P*=.25), relationship with caregiver (*P*=.57), and dosage outcome (*P*=.77).

Figure 8. Flow diagram of the participants' progress through the study phases. CG: control group; MCI: mild cognitive impairment; TG: treatment group.

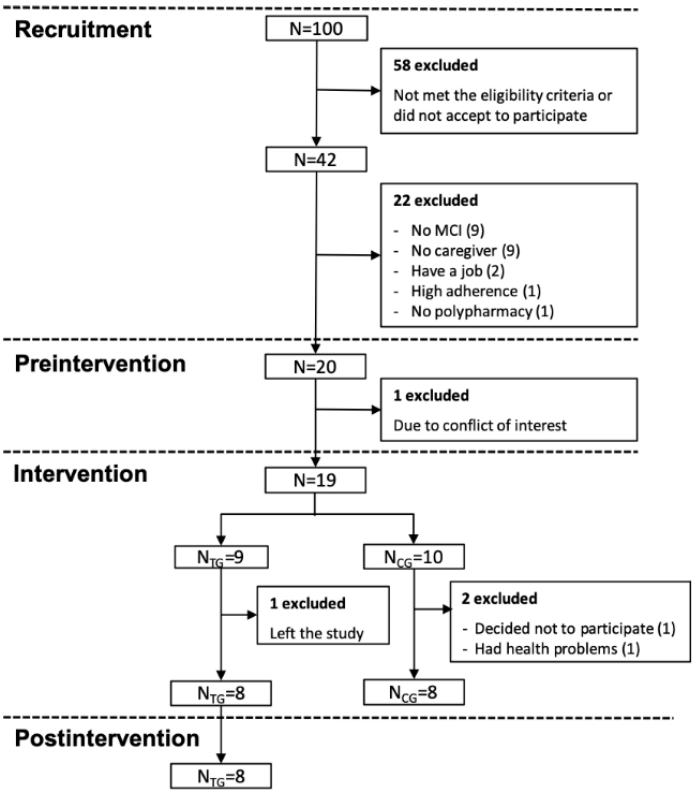


Table 3. Characteristics of the participants.

Characteristic	Group		Statistics		
	Control	Treatment	*t* test (*df*)	Chi-square test (*df*)	*P* value
Age (years), mean (SD)	73.5 (8.3)	68.62 (6.2)	1.32 (15)	—[a]	.21
Education (years), mean (SD)	5.25 (3.8)	6.75 (2.1)	0.97 (15)	—	.35
Number of medications, mean (SD)	5.75 (1.8)	4.88 (1.6)	1.01 (15)	—	.33
Cognition	Mild[b]	Mild[b]	—	—	—
Sex, n				1.3 (1)	.25
Female	5	7	—		
Male	3	1	—		
Caregiver relationship, n				1.1 (1)	.57
Spouse	3	3	—		
Child	4	5	—		
Other	1	0	—		
Self-reported medication adherence, n				0.2 (1)	.59
Low	2	3	—		
Middle	6	5	—		
Pill counting medication adherence (%), mean (SD)				0.30 (15)	.77
Baseline	79.87 (17.9)	80.9 (16)	—		

[a]Not applicable.

[b]Cognitive impairment was assessed with the Short Portable Mental Status Questionnaire because it is appropriate for low-literacy persons [26].

Adherence to Medication (Research Question 1)

Although underadherence was the most predominant behavior, overadherence was also found during the stages of different studies, for both the TG and the CG (see Multimedia Appendix 1). We identified overmedication and undermedication during the intervention phase not only according to the pill counting technique but also according to the medication episodes recorded in the MAD's log. For instance, Multimedia Appendix 1 shows that participant P2 has adherence rates higher than 100% according to both techniques. Similarly, overmedication could be present in the baseline data, which impacted the high medication rates registered for some older adults (eg, participants P7 and P8). However, our results show that the introduction of the MAD's external cues to the TG resulted in significant improvements in the average rates of dosage outcomes, which addresses RQ1.

$Dosage_{pill}$ (Treatment Group)

The TG improved their adherence behavior (dosage), increasing from 80.9% in the preintervention phase to 95.97% in the intervention phase. However, it decreased to 76.71% in the postintervention phase. Using a one-way repeated measures ANOVA, we compared the effect of the MAD in the TG during the 3 phases. It showed a significant statistical difference between at least two of the phases ($F_{2,14}$=6.59; P=.0096). With a post hoc analysis using the Tukey honestly significant difference test, we identified a statistical difference between the preintervention and intervention phases (P=.02) and between the intervention and postintervention phases (P=.002); Cohen effect size values (d=1.35 and d=1.72, respectively) suggest a high practical significance in both cases. In addition, there was no statistical difference between the preintervention and the postintervention phases (P=.73). Therefore, these results are evidence that the external cues of the MAD contributed to improving the medication intake behaviors of older adults.

$Dosage_{pill}$ (Control Group)

The CG increased their medication adherence between the preintervention (mean 79.87% [SD 20.49%]) and intervention phases (mean 88.18% [SD 20.49%]). However, according to a paired-samples t test, there was not a significant difference (t_7=1.15; P=.14). The Cohen effect size value (d=0.41) suggests a medium practical significance. We attributed this increase in dosage rates to the medication monitoring conducted weekly by the research assistants and to the basket provided to organize the pills containers.

$Dosage_{pill}$ (Treatment Group vs Control Group)

The medication adherence rate of the TG (mean 95.97% [SD 6.08%]) was higher than that of the CG (mean 88.18% [SD 13.06%]) during the intervention phase. However, according to a independent samples t test, there was not a significant difference (t_{14}=1.53; P=.08). The Cohen effect size value (d=0.76) suggests a high practical significance. The monitoring conducted by research assistants impacted both groups.

$Dosage_{pill}$ Versus $Dosage_{MAD}$ (Treatment Group)

A one-way ANOVA was conducted to compare the effect of $Dosage_{pill}$ versus $Dosage_{MAD}$ on the adherence of the TG. An ANOVA showed that the effect of both measuring instruments on adherence was not significant ($F_{1,14}$=0.21; P=.65). This confirms the suitability of the results obtained through the pill counting technique to explain how the MAD supports adherence to medication. In addition, as described in the following results, $Dosage_{MAD}$ estimation enabled us to determine how the MAD reminders helped older adults take their medications at the prescribed times.

Timely and Reminder Dependency Rates

We obtained 2224 medication episodes registered in the MAD's log, of which 93.17% (2072/2224) were registered as taken on time, and 88.35% (1830/2224) of these timely episodes were taken after the MAD reminders. As illustrated in Table 4, most of the participants showed a high reminder dependency rate. These results are evidence that the MAD reminders resulted in medication-taking behaviors consistently.

Table 4. Timely and reminder dependency rates estimated by participants.

Participants in the treatment group	Medication episodes rate (%)	Timely episodes rate (%)	Reminder dependency rate (%)
P1	86.46	87.37	93.64
P2	100.67	92	94.2
P3	93.83	94.92	100
P4	100	92.44	32.19
P5	94.71	97.21	98.56
P6	99.23	89.92	93.97
P7	98.56	98.06	99.5
P8	100.43	93.16	98.62

Self-Reported Medication Adherence

We found that MAQ-8 scores and $Dosage_{pill}$ outcomes are not correlated but are independent of each other (N=16; ρ=−0.29; P=.27). A McNemar test determined that there was no statistically significant difference in the MAQ-8 scores obtained in the preintervention and postintervention stages (P=.25). These results suggest that using the MAD did not change older adults' perception about their medication adherence.

System Adoption (Research Question 2)

The collected qualitative evidence helps us to address RQ2 and to complement the quantitative results. We analyzed the data collected from the interviews administered to participants P1 to P8 of the TG, and we discovered the findings described in the following subsections.

Usability of the Ambient Modalities That Implement the Medication Ambient Display's External Cues

All the older adults provided answers that lead us to conclude that the most useful cues were the auditory reminders, followed by the stylized representation of medication adherence (eg, parakeet growth). Some participants explained how these external cues caught their attention so that they can medicate themselves. Participant P6 said:

> *Sometimes I am busy or just thinking about something else, and I forgot what I have to do; but now when the parakeet sings and sings, it reminds me to medicate.*

This finding is supported by the high rates in timely and reminder dependency measures. For instance, participant P1 improved her medication-taking behavior because she developed a high dependency on the ambient reminders (see Table 4), which affected her medication compliance during the postintervention phase (see Multimedia Appendix 1). She reported:

> *When I heard the parakeet whistle, I came to the kitchen to take my medicines...It was better when I had the system.*

All older adults, except participant P4, reported that they did not consult the detailed information on their medication adherence, but they verified if the parakeet grew after registering the medication as taken. Thus, participant P4 was the only older adult who did not develop a high reminder dependency rate (see Table 4) to stabilize her medication adherence (see Multimedia Appendix 1), but she relied on the stylized abstract representations and detailed report of her adherence; in addition, she expressed interest in consulting her medication adherence from the last week or month.

Older Adults' Perception of Their Medication Adherence

Although adverse drug events associated with overmedication and undermedication were identified, older adults hardly recognized that before using the MAD; there were times when they might have forgotten to take a medication. For instance, although participant P4 explicitly stated that she did not forget to take her pills, her husband contradicted her. Participant P6 was the only participant who, during the interviews, admitted forgetting to take her medication because as a consequence, she had symptoms of her disease:

> *[Before using the MAD] I realized that I had forgotten to take the [night doses] pill for my blood pressure, until the next morning my head hurt and I felt dizzy.*

On the other hand, we have no evidence that older adults were aware that they sometimes overmedicate because they forgot that they had taken their medication.

Integration of Medication Routine Into Daily Activities

Older adults intend to take their medications when carrying out some of their daily routines. We identified that some older adults fail to associate medication with their daily activities effectively. For instance, participant P3 indicated:

> *I plan to take my night medication before going to sleep, but before using MAD, sometimes I realized that I had forgotten to medicate until I was in bed, and murmured: "Ay! I have not taken the pill." However, now, it [MAD] sounds, I take my medication, and then, I go to sleep.*

On the other hand, we found that leaving home or traveling affected the use of the MAD and probably the medication adherence of participants. For instance, several participants (P2, P3, and P6) reported situations in which they could not register a medication intake timely because they had to leave home. For instance, participant P6 reported:

> *I did not take my medication this Wednesday since I had to leave the house urgently, but I took it later.*

Also, only 1 participant (P4) felt confident enough to take the system with her when she left home.

External Cues Mediated Caregivers' Assistance

The family caregivers commonly served as the main support actor and assisted with managing the seniors' medications. Before using the MAD, caregivers tended to be aware of the medication timetables to remind older adults or ask older adults if they had medicated. We observed that the external cues of the MAD acted as triggers that facilitated caregivers' assistance. That is, the cues did not overwhelm family members but were an appropriate mediation strategy to support seniors' medication routines. For instance, the husband of participant P4 perceived that the MAD enabled him to be aware when she took the correct medications. For an adolescent caregiver, the system enabled him to feel less worried about having to remind her grandmother (participant P1) to medicate:

> *If I have to do my homework, I can focus on doing it.*

The MAD also helped to assure caregivers that older adults would not forget to take their medications; for example, participant P6 said:

> *My children used to forget reminding me to medicate [before using the MAD]. Currently, they hear the parakeet, and then they make sure if I took them.*

Discussion

Principal Findings

Our quantitative results show that providing the external cues supported by the MAD resulted in significant improvements in the average rates of dosage outcomes for older adults. This is because the ambient modalities used for implementing these external cues were useful for drawing the attention of older adults. We found that external cues (1) reminded them to take medications, (2) enabled them to recognize if a medication was recorded as taken, and (3) provided immediate awareness about how they followed their medication regimens.

We learned that providing older adults with an abstract and stylized representation of their medication adherence, which could be peripherally perceived, was better accepted than medication adherence reports that need to be explicitly evoked. However, this stylized and abstract modality of representation was not enough to make participants aware of their medication problems related to undermedication and overmedication. Previous research has demonstrated that feedback-based systems that are consulted explicitly and daily help seniors identify their medication errors and then self-regulate their medication behavior [20]. We consider that providing timely detailed information about medication adherence may help older adults perceive the usefulness of systems designed to aid medication uptake and, therefore, encourage their adoption. From our study, we learned that it is necessary to make the feedback more salient to provide them with sufficient knowledge of their medication problems.

Similar to our results, other studies have shown that when older adults stop using medication aiding systems, their medication adherence is affected [20]. Our participants showed high dependence on the MAD reminders; therefore, when the system was removed from their homes, their adherence to medications was negatively affected. We also found that older adults tend to link their medication to other daily tasks [1,37], for example, take medication before going to sleep. We conclude that ambient displays should be flexible enough to adapt external cues according to the activities that older adults associate with their medication intake behaviors. These include taking into account the daily routines in which the medication should be inserted (eg, before meals), in addition to taking into account context changes (eg, going out from home). We hypothesized that including adaptation mechanisms to the MAD that allow seniors to configure external cues to remind them to take medications in a specific context instead of a specific window of time would allow that medication to be integrated into their daily routine. For instance, if an older adult proposes to take their medications when preparing his or her coffee in the morning, the MAD could include cues to help them remember to associate their medication with that activity. In this case, the MAD would be training older adults during a period to take their medication when that specific context arises. We hypothesized that supporting this strategy may reduce the dependency of older adults to the MAD's medication reminders, as it may help them develop the habit of taking their medication in the same context consistently.

Furthermore, we identified that external cues of the MAD provided caregivers with better awareness of older adults' medication adherence. This awareness was 2-fold: (1) when auditory reminders were perceived by caregivers, they made sure that the reminders reached the target recipients, and (2) medication adherence representations enabled caregivers to be aware of medications taken. Therefore, providing external cues through ambient displays helps family caregivers to better support seniors to follow their medication regimens.

Limitations

Economic Incentive for Participating

The use of financial incentives has been questioned because they may provide inducements to participate in a study for financial purposes only, and vulnerable populations are prone to be enticed by the financial reward and be more willing to accept any study risks [38]. In our opinion, offering an incentive facilitated recruitment of participants and allowed us to access their data, which would otherwise be considered an obtrusive task. For instance, 2 participants explicitly asked for the weekly economic incentive to allow the research assistants to collect data; another participant questioned the incentive since she considered it was unnecessary for being part of the study. Although we are not able to conclude on how the incentive impacted the adoption of the technology, our findings indicate that the MAD was accepted not only by older adults but also by their family caregivers.

Setting Characteristics

We observed that some participants faced problems in managing their medications, such as accumulating medications, confusing medications because they look alike, and tending to give medications to others, which may not be an appropriate practice. The design of our study was limited in that the MAD system was personalized according to the prescribed medication regimens and timetables that the participants followed. Although using our technology did not introduce any risk, it might have supported inappropriate medication routines adopted by older adults to overcome some of the barriers imposed by the setting. We recognize the importance of conducting a contextual study before conducting a technology evaluation. The contextual study should be designed in collaboration with clinical or nursing specialists to reduce the complexity of older adults' medication regimens and the risks associated with the way they manage medications. Finally, we are not able to state that our results are generalizable to the whole Mexican elderly population. This is because participants were primarily from a low-income socioeconomic stratum; therefore, exploring this technology among high-income elders who have access to private health care services might produce different findings.

Study Methodology

The pill counting technique can be prone to human error. So, one limitation of this study is that we were not able to identify which overmedication and undermedication events were registered by the research assistants erroneously. Using electronic monitoring devices (EMDs), such as the Medication Event Monitoring System, could have overcome this limitation to some extent, although using EMDs does not guarantee that a person has taken their medication [39]. Moreover, the cost of EMDs and the logistics of integrating them within clinical protocols may be limiting factors to their adoption for research in clinical contexts [39]. One possible solution to overcome this limitation could be to combine data collection techniques.

Study Duration

Another limitation of the study is its duration time. Some research works have identified that older adults should have an adaptation period to an intervention, and after a specific period

of using it, reliable data can be collected to measure its effectiveness for improving medication compliance [3]. We provided evidence that our study enabled us to get an understanding of the feasibility of the MAD to improve adherence to medication. However, extending the time of both stages, intervention and postintervention, would have allowed us to understand the efficacy of the MAD to sustain high rates of medication adherence.

Comparison With Prior Studies

Table 5 shows an overview of some studies published in the last decade, which were conducted to assess technological-based interventions to support behavioral strategies to improve older adults' medication adherence, such as providing reminders [18,19,40-43], medication feedback [20,39-41], and self-monitoring of pills [20,39,42]. Some of these studies also included a patient education strategy [15,42,43], which is considered a traditional approach used to tackle the medication adherence problem [44]. The technologies evaluated in these studies included mobile phones, tablets, EMDs, and pill dispensers. These studies have assessed the technologies' effect on medication adherence, in addition to evaluating older adults' acceptance. Most of them used both subjective (eg, self-reporting) and objective (eg, pill counting and system logs) adherence assessment methods [15,20,40-43] because no single method is sufficiently reliable and accurate [30]. Nonetheless, some of these studies are limited to assessing only the adherence to medications taken for a particular health problem [40,41] and did not take into account older adults with multiple morbidities and polypharmacy [15,39,40], which are factors that contribute to increasing the risk of nonadherence [45,46]. Other reasons for nonadherence include low literacy, cultural factors, and inadequate social support [45]. In this sense, none of the studies presented in Table 5 examined the involvement of family caregivers, although some of them report that their technologies offered mechanisms to enable the family caregivers' participation [19,39-41]. In contrast, our study assessed the effect of our approach by using objective medication adherence measures; in addition, we obtained qualitative findings that explained the adoption of the MAD from the perspectives of both seniors and their relatives.

Table 5. Overview of some studies to assess different technological approaches to support older adults' medication adherence.

Data extracted from the studies	Moraws-ki et al [40]	Mertens et al [41]	Robiner et al [39]	Grindrod et al [18]	Park et al [15]	Lee and Dey [20]	Perera et al [42]	Patel et al [43]	Reeder et al [19]	De Oliveira et al [21]
Technology										
Mobile phone apps	X[a]	—[b]	—	—	—	—	X	X	—	X
SMS	—	—	—	—	X	—	—	—	—	—
Tablet	—	X	—	X	—	X	—	—	—	—
Dispenser	—	—	—	—	—	—	—	—	X	—
Monitoring	—	—	X	—	—	X	—	—	—	—
Functions										
Remind	X	X	—	X	—	—	X	X	X	—
Register taken doses	X	X	—	—	—	—	X	X	X	X
Educate	—	—	—	—	X	—	X	X	—	—
Feedback	X	X	X	—	—	X	—	—	—	—
Caregiver participation	X	X	X	—	—	—	—	—	X	—
Social game	—	—	—	—	—	—	—	—	—	X
Evaluation										
Total participants (older adults)	413	24	6	35	90	12	28	48	96	16
Medicines	$2_{\text{Hiperten-sion}}$	$+3_{\text{cardvas-cular de-sease}}$	1_{renal}	$+1_{\text{several}}$	$2_{\text{cardiovascu-lar desease}}$	N/S[c]	3_{HIV}	$3_{\text{Hiperten-sion}}$	$+11_{\text{several}}$	$+1_{\text{serveral}}$
Methods to measure adherence: subjective and objective	Subjective and objective	Subjective and objective	Objective	—	Subjective and objective	Subjective and objective	Subjective and objective	Subjective and objective	—	—
Acceptance	—	X	X	X	X	—	X	X	X	X

[a]X: studies that assessed the acceptance of the system by the participants.

[b]Not applicable.

[c]N/S: nonsignificant.

Conclusions

The external cues provided by our ambient display not only improved the medication adherence of the elderly but also encouraged caregiver involvement. We found that the external cues perceived as most useful were those that reminded participants to take medications, helped seniors recognize if medications were recorded as taken, and provided immediate and abstract representations of their medication adherence. We identified that external cues did not overwhelm family members but were an appropriate mediation strategy to support older adults' medication routines. We also recognized the potential of providing external cues to enable older adults to associate medication routines with their daily routines appropriately. For future work, we plan to conduct studies to assess the feasibility of external ambient cues to support the seamless integration of medication regimens into the daily routines of the elderly.

Acknowledgments

The National Council of Science and Technology (CONACyT) in Mexico and the Autonomous University of Baja California supported this work under grant numbers 153863 and 1914, respectively. The authors want to thank CONACyT for the scholarships provided to the first and seventh authors. The authors thank the students from the Faculty of Nursing who helped with participant

recruitment and data collection, and professors Betzabé Arizona, Rosa Esparza, and José Aguero for their support with conducting the study.

Conflicts of Interest

None declared.

This randomized study was only retrospectively registered, explained by authors as: “our paper presents a small study to provide evidence about the feasibility to be accepted and its efficacy to improve the medication adherence of older adults.” The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low and the study was considered formative, guiding the development of the application. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
Individual dosage rates estimated through the pill counting technique (Dosagepill) and the medication episodes recorded in MAD’s log (DosageMAD). MAD: Medication Ambient Display.
[PNG File , 97 KB - mhealth_v8i3e14680_app1.png]

Multimedia Appendix 2
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1386 KB - mhealth_v8i3e14680_app2.pdf]

References

1. Insel KC, Einstein GO, Morrow DG, Hepworth JT. A multifaceted prospective memory intervention to improve medication adherence: design of a randomized control trial. Contemp Clin Trials 2013 Jan;34(1):45-52 [FREE Full text] [doi: 10.1016/j.cct.2012.09.005] [Medline: 23010608]
2. Murray MD, Morrow DG, Weiner M, Clark DO, Tu W, Deer MM, et al. A conceptual framework to study medication adherence in older adults. Am J Geriatr Pharmacother 2004 Mar;2(1):36-43. [Medline: 15555477]
3. Insel KC, Einstein GO, Morrow DG, Koerner KM, Hepworth JT. Multifaceted prospective memory intervention to improve medication adherence. J Am Geriatr Soc 2016 Mar;64(3):561-568 [FREE Full text] [doi: 10.1111/jgs.14032] [Medline: 27000329]
4. Gillespie R, Mullan J, Harrison L. Managing medications: the role of informal caregivers of older adults and people living with dementia. A review of the literature. J Clin Nurs 2014 Dec;23(23-24):3296-3308. [doi: 10.1111/jocn.12519] [Medline: 24354583]
5. Kang HS, Myung W, Na DL, Kim SY, Lee J, Han S, et al. Factors associated with caregiver burden in patients with Alzheimer's disease. Psychiatry Investig 2014 Apr;11(2):152-159 [FREE Full text] [doi: 10.4306/pi.2014.11.2.152] [Medline: 24843370]
6. Smith G, Della Sala S, Logie RH, Maylor EA. Prospective and retrospective memory in normal ageing and dementia: a questionnaire study. Memory 2000 Sep;8(5):311-321. [doi: 10.1080/09658210050117735] [Medline: 11045239]
7. Gillespie RJ, Harrison L, Mullan J. Medication management concerns of ethnic minority family caregivers of people living with dementia. Dementia (London) 2015 Jan;14(1):47-62. [doi: 10.1177/1471301213488900] [Medline: 24339089]
8. Lingler JH, Sereika SM, Amspaugh CM, Arida JA, Happ ME, Houze MP, et al. An intervention to maximize medication management by caregivers of persons with memory loss: Intervention overview and two-month outcomes. Geriatr Nurs 2016;37(3):186-191 [FREE Full text] [doi: 10.1016/j.gerinurse.2015.12.002] [Medline: 26804450]
9. While C, Duane F, Beanland C, Koch S. Medication management: the perspectives of people with dementia and family carers. Dementia (London) 2013 Nov;12(6):734-750. [doi: 10.1177/1471301212444056] [Medline: 24337637]
10. Cramer JA. Enhancing patient compliance in the elderly. Role of packaging aids and monitoring. Drugs Aging 1998 Jan;12(1):7-15. [doi: 10.2165/00002512-199812010-00002] [Medline: 9467683]
11. García-Vázquez JP, Rodríguez MD, Andrade AG, Bravo J. Supporting the strategies to improve elders’ medication compliance by providing ambient aids. Pers Ubiquit Comput 2011;15(4):389-397. [doi: 10.1007/s00779-010-0362-0]
12. Rodríguez M, Zárate E, Stawarz K, García-Vázquez J, Ibarra E. Ambient computing to support the association of contextual cues with medication taking. Rev Mex Ing Bioméd 2015;36(3):189-205. [doi: 10.17488/RMIB.36.3.5]
13. Zárate-Bravo E, García-Vázquez JP, Rodríguez MD. An ambient medication display to heighten the peace of mind of family caregivers of older adults: A study of feasibility. In: Serino S, Matic A, Giakoumis D, Lopez G, Cipresso P, editors. Pervasive Computing Paradigms for Mental Health. MindCare 2015. Communications in Computer and Information Science. Volume 604. Switzerland: Springer International Publishing; 2016:274-283.

14. Cramer JA, Roy A, Burrell A, Fairchild CJ, Fuldeore MJ, Ollendorf DA, et al. Medication compliance and persistence: terminology and definitions. Value Health 2008;11(1):44-47 [FREE Full text] [doi: 10.1111/j.1524-4733.2007.00213.x] [Medline: 18237359]
15. Park LG, Howie-Esquivel J, Chung ML, Dracup K. A text messaging intervention to promote medication adherence for patients with coronary heart disease: a randomized controlled trial. Patient Educ Couns 2014 Feb;94(2):261-268. [doi: 10.1016/j.pec.2013.10.027] [Medline: 24321403]
16. Park LG, Howie-Esquivel J, Dracup K. A quantitative systematic review of the efficacy of mobile phone interventions to improve medication adherence. J Adv Nurs 2014 Sep;70(9):1932-1953. [doi: 10.1111/jan.12400] [Medline: 24689978]
17. Stawarz K, Cox L, Blandford A. Don't Forget Your Pill!: Designing Effective Medication Reminder Apps That Support Users' Daily Routines. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI'14; April 26-May 1, 2014; Toronto, Ontario, Canada p. 2269-2278. [doi: 10.1145/2556288.2557079]
18. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014 Mar 14;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]
19. Reeder B, Demiris G, Marek KD. Older adults' satisfaction with a medication dispensing device in home care. Inform Health Soc Care 2013 Sep;38(3):211-222 [FREE Full text] [doi: 10.3109/17538157.2012.741084] [Medline: 23323721]
20. Lee M, Dey A. Real-Time Feedback for Improving Medication Taking. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI'14; April 26-May 1, 2014; Toronto, Ontario, Canada p. 2259-2268 URL: https://dl.acm.org/citation.cfm?doid=2556288.2557210 [doi: 10.1145/2556288.2557210]
21. de Oliveira R, Cherubini M, Oliver N. MoviPill: Improving Medication Compliance for Elders Using a Mobile Persuasive Social Game. In: Proceedings of the 12th ACM international conference on Ubiquitous computing. 2010 Presented at: UbiComp'10; September 26-29, 2010; Copenhagen, Denmark p. 251-260 URL: https://dl.acm.org/citation.cfm?doid=1864349.1864371 [doi: 10.1145/1864349.1864371]
22. Tschanz M, Dorner T, Holm J, Denecke K. Using eMMA to manage medication. Computer 2018;51(8):18-25 [FREE Full text] [doi: 10.1109/MC.2018.3191254]
23. Lobo J, Ferreira L, Ferreira A. CARMIE: A conversational medication assistant for heart failure. J E-Health Med Commun 2017;8(4):21-37 [FREE Full text] [doi: 10.4018/ijehmc.2017100102]
24. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 6;10:1 [FREE Full text] [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
25. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
26. Iglesiaa J, Dueñas R, Vilchesa M, Tabernéa C, Colomerc C, Luquec R. Adaptación y validación al castellano del cuestionario de Pfeiffer (SPMSQ) para detectar la existencia de deterioro cognitivo en personas mayores e 65 años. Medicina Clínica 2001;117(4):129-134. [doi: 10.1016/S0025-7753(01)72040-4]
27. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986 Jan;24(1):67-74. [doi: 10.1097/00005650-198601000-00007] [Medline: 3945130]
28. Lavsa SM, Holzworth A, Ansani NT. Selection of a validated scale for measuring medication adherence. J Am Pharm Assoc (2003) 2011;51(1):90-94. [doi: 10.1331/JAPhA.2011.09154] [Medline: 21247831]
29. Orwig D, Brandt N, Gruber-Baldini AL. Medication management assessment for older adults in the community. Gerontologist 2006 Oct;46(5):661-668. [doi: 10.1093/geront/46.5.661] [Medline: 17050757]
30. Giardini A, Martin MT, Cahir C, Lehane E, Menditto E, Strano M, et al. Toward appropriate criteria in medication adherence assessment in older persons: Position Paper. Aging Clin Exp Res 2016 Jun;28(3):371-381. [doi: 10.1007/s40520-015-0435-z] [Medline: 26630945]
31. Wirtz VJ, Reich MR, Flores RL, Dreser A. Medicines in Mexico, 1990-2004: systematic review of research on access and use. Salud Publica Mex 2008;50(Suppl 4):S470-S479. [doi: 10.1590/s0036-36342008001000008] [Medline: 19082258]
32. Mahmood A, Elnour AA, Ali AA, Hassan NA, Shehab A, Bhagavathula AS. Evaluation of rational use of medicines (RUM) in four government hospitals in UAE. Saudi Pharm J 2016 Mar;24(2):189-196 [FREE Full text] [doi: 10.1016/j.jsps.2015.03.003] [Medline: 27013912]
33. Mao W, Vu H, Xie Z, Chen W, Tang S. Systematic review on irrational use of medicines in China and Vietnam. PLoS One 2015;10(3):e0117710 [FREE Full text] [doi: 10.1371/journal.pone.0117710] [Medline: 25793497]
34. 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]
35. Montgomery DC. Design And Analysis Of Experiments. Eighth Edition. North Caroline, USA: Wiley; 2012.
36. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
37. Stawarz K, Rodríguez MD, Cox AL, Blandford A. Understanding the use of contextual cues: design implications for medication adherence technologies that support remembering. Digit Health 2016;2:2055207616678707 [FREE Full text] [doi: 10.1177/2055207616678707] [Medline: 29942574]

38. Groth SW. Honorarium or coercion: use of incentives for participants in clinical research. J N Y State Nurses Assoc 2010;41(1):11-3; quiz 22 [FREE Full text] [Medline: 20882821]
39. Robiner WN, Flaherty N, Fossum TA, Nevins TE. Desirability and feasibility of wireless electronic monitoring of medications in clinical trials. Transl Behav Med 2015 Sep;5(3):285-293 [FREE Full text] [doi: 10.1007/s13142-015-0316-1] [Medline: 26327934]
40. Morawski K, Ghazinouri R, Krumme A, McDonough J, Durfee E, Oley L, et al. Rationale and design of the Medication adherence Improvement Support App For Engagement-Blood Pressure (MedISAFE-BP) trial. Am Heart J 2017 Apr;186:40-47. [doi: 10.1016/j.ahj.2016.11.007] [Medline: 28454831]
41. Mertens A, Brandl C, Miron-Shatz T, Schlick C, Neumann T, Kribben A, et al. A mobile application improves therapy-adherence rates in elderly patients undergoing rehabilitation: A crossover design study comparing documentation via iPad with paper-based control. Medicine (Baltimore) 2016 Sep;95(36):e4446 [FREE Full text] [doi: 10.1097/MD.0000000000004446] [Medline: 27603339]
42. Perera AI, Thomas MG, Moore JO, Faasse K, Petrie KJ. Effect of a smartphone application incorporating personalized health-related imagery on adherence to antiretroviral therapy: a randomized clinical trial. AIDS Patient Care STDS 2014 Nov;28(11):579-586 [FREE Full text] [doi: 10.1089/apc.2014.0156] [Medline: 25290556]
43. 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 1;7(3):630-639 [FREE Full text] [doi: 10.1177/193229681300700307] [Medline: 23759395]
44. Banning M. A review of interventions used to improve adherence to medication in older people. Int J Nurs Stud 2009 Nov;46(11):1505-1515. [doi: 10.1016/j.ijnurstu.2009.03.011] [Medline: 19394018]
45. Williams A, Manias E, Walker R. Interventions to improve medication adherence in people with multiple chronic conditions: a systematic review. J Adv Nurs 2008 Jul;63(2):132-143. [doi: 10.1111/j.1365-2648.2008.04656.x] [Medline: 18537843]
46. Claxton AJ, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. Clin Ther 2001 Aug;23(8):1296-1310. [doi: 10.1016/s0149-2918(01)80109-0] [Medline: 11558866]

Abbreviations

ANOVA: analysis of variance
CG: control group
CONACyT: National Council of Science and Technology
EMD: electronic monitoring device
IMSS: Mexican Institute of Social Security
MAD: Medication Ambient Display
MAQ-8: 8-item Medication Adherence Questionnaire
MedMaIDE: Medication Management Instrument for Deficiencies in the Elderly
NFC: Near Field Communication
RQ: research question
TG: treatment group
UNICOM: Community Center of the University

Edited by G Eysenbach; submitted 21.05.19; peer-reviewed by R Hervás, A Carr; comments to author 22.06.19; revised version received 10.09.19; accepted 24.09.19; published 02.03.20.

Please cite as:
Zárate-Bravo E, García-Vázquez JP, Torres-Cervantes E, Ponce G, Andrade ÁG, Valenzuela-Beltrán M, Rodríguez MD
Supporting the Medication Adherence of Older Mexican Adults Through External Cues Provided With Ambient Displays: Feasibility Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(3):e14680
URL: http://mhealth.jmir.org/2020/3/e14680/
doi:10.2196/14680
PMID:32130164

Original Paper

Preliminary Effects of a Mobile Interactive Supervised Therapy Intervention on People Living With HIV: Pilot Randomized Controlled Trial

Yan Pang[1,2], MSc; James Steven Molton[2,3], MBBS; Wei Tsang Ooi[4], BSc, PhD; Nicholas Iain Paton[2,3], MD, FRCP; Hong-Gu He[1], MSc, PhD

[1]Alice Lee Centre for Nursing Studies, National University of Singapore, Singapore, Singapore

[2]University Medicine Cluster, National University Health System, Singapore, Singapore

[3]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[4]School of Computing, National University of Singapore, Singapore, Singapore

Corresponding Author:
Yan Pang, MSc
Alice Lee Centre for Nursing Studies
National University of Singapore
NUHS Tower Block Level 8, 1E Kent Ridge Road
Singapore, 119228
Singapore
Phone: 65 60015375
Email: e0011744@u.nus.edu

Abstract

Background: As people living with HIV infection require lifelong treatment, nonadherence to medication will reduce their chance of maintaining viral suppression and increase the risk of developing drug resistance and HIV transmission.

Objective: This study aimed to evaluate the efficacy of a mobile app, Mobile Interactive Supervised Therapy (MIST), for improving adherence to oral HIV medications among HIV-infected adults in Singapore.

Methods: We conducted a two-group pilot randomized controlled trial (RCT) with a process evaluation, in which 40 HIV-infected participants with once-daily medication regimes were recruited from a public tertiary hospital in Singapore and randomly assigned equally to either the intervention (receiving MIST and routine care) or control (receiving routine care only) groups. The intervention lasted for 2 months. The outcome of antiretroviral therapy (ART) adherence was measured by a 7-day recall self-report (SR), pill count (PC), an electronic medical device—Medication Event Monitoring System (MEMS)—and a mobile app—MIST (for the intervention group only). In total, 20 participants from the intervention group were interviewed at the end of the intervention to assess the acceptability of MIST. Data were collected at baseline and at 1-month and 2-month postintervention.

Results: All participants had excellent medication adherence at baseline (median 100, IQR 100-100). The use of MIST did not result in a significant improvement in ART adherence when measured by the SR, PC, and MEMS, as compared with the control group at 1-month (*P* values >.99, .86, and .74, respectively) and 2-month (*P* values=.80, .84, and .82, respectively) postintervention. ART adherence also did not improve in each group over the same period. MIST was perceived to be a beneficial tool based on the process evaluation results.

Conclusions: Although MIST did not enhance medication adherence to HIV treatments, mainly owing to the ceiling effect, it was perceived to be beneficial among the participants of this study. Our process evaluation provided useful data to further develop MIST for bigger and long-term mobile phone app–assisted intervention RCTs in the future.

Trial Registration: ClinicalTrials.gov NCT03794648; https://clinicaltrials.gov/ct2/show/NCT03794648

(JMIR Mhealth Uhealth 2020;8(3):e15702) doi:10.2196/15702

KEYWORDS

antiretroviral treatment, highly active; human immunodeficiency virus; medication adherence; mobile application

Introduction

Background

HIV infection is a global public health issue. An estimated 36.9 million people were living with HIV in 2017 [1]. In the same year, 7982 residents of Singapore were living with HIV, with 434 being new cases [2]. In people living with HIV, who require lifelong treatment, nonadherence to medication not only reduces their chance of maintaining viral suppression but also increases the risk of developing drug resistance and HIV transmission [3-6]. To achieve viral suppression, a patient needs to maintain an adherence level of at least 90% throughout the treatment period [6-8]. However, there are challenges to meeting this target. A meta-analysis showed that only 62% of patients of a pooled study aged older than 18 years taking prescribed highly active antiretroviral therapy reported an adherence rate of ≥90% [9]. There are no available antiretroviral therapy (ART) adherence data in Singapore.

With phone technology evolving, mobile phones and smartphones are now able to incorporate many creative features such as recording and transmitting high-quality videos and telemedicine capabilities. There have been a growing number of mobile app reminders supporting patients' adherence to medication in the market; however, there is a lack of robust evidence to establish their efficacy. To date, the quality of these supporting systems has been variable [10,11]. In addition, current evidence supporting the effectiveness of mobile phone app–based interventions in enhancing adherence among people living with HIV is unclear because of limited evidence and the heterogeneity of the study populations, treatments, and ART adherence measurements [12-17].

Aims

This study aimed to examine the effects of a mobile phone app, Mobile Interactive Supervised Therapy (MIST), for improving adherence to oral HIV medications among HIV-infected adult patients in Singapore and to gather user experiences to develop the intervention for large scale randomized controlled trials (RCTs) in the future.

Hypotheses

The hypotheses were as follows:

1. When compared with the control group, patients in the MIST intervention group will have better HIV medication adherence rates at the 1-month and 2-month follow-up time points.
2. When compared with baseline, patients in the MIST intervention group will have better HIV medication adherence rates at the end of the study.

Methods

Design

A two-group pretest and posttest RCT design was used. Adult patients infected with HIV (n=40) were recruited from a public hospital in Singapore. Recruited participants were randomly assigned into either of the two groups: the intervention group or the control group.

Participants

All adults infected with HIV who attended their routine clinic appointments in a public hospital in Singapore from March to June 2018 were approached to participate in this study. The inclusion criteria included those who were (1) aged ≥21 years, (2) taking a once daily regimen of HIV medications, (3) able to take pills orally, (4) willing and able to give informed consent, and (5) able to read and speak English or Chinese. Participants were excluded if they (1) were unable to operate a mobile phone or had an active tuberculosis infection that required directly observed therapy during the study period, (2) had substance use such as methamphetamine use, (3) had visual, speech, or hearing impairment despite the use of aids, (4) had a known medical history of psychiatric disorder(s) and were seeking any form of psychiatric treatment, (5) had a known medical history of cognitive impairment, (6) were fully dependent on a caregiver for taking medications, (7) were pregnant at the time of recruitment and data collection, (8) had a terminal illness such as cancer or late stage cardiovascular disease, or (9) had experienced bereavement within the past 6 months.

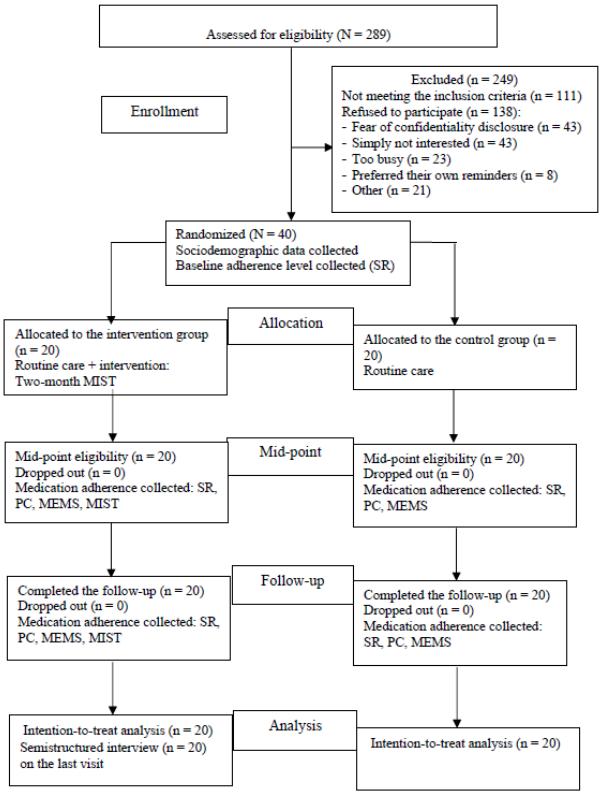
Sample Size Determination

On the basis of pragmatic considerations of the pilot study, a sample of 40 participants, with 20 in each group, was decided a priori [18-20]. All 20 participants in the intervention group were invited for the final process evaluation. There was no dropout during the follow-up period.

Randomization

After obtaining written consent, baseline sociodemographic (age, gender, ethnicity, and education level) and clinical data (ART regimen and duration), phone operation technical skills, and a 7-day recall of HIV medication adherence data were collected, and participants were randomly assigned to two groups: the intervention group (receiving the MIST intervention plus routine care) or the control group (receiving routine care only) in a 1:1 allocation ratio. Block randomization with *length 4* was used [21] to ensure a balanced representation in each group [22]. Participants were each asked to open an opaque, sealed envelope with a piece of cardboard inside indicating their randomly allocated groups and were assigned in a successive order according to their enrollment sequence. Through these processes, randomization and allocation concealments were ensured [23-25]. The detailed workflow is described using the Consolidated Standards of Reporting Trial (CONSORT) flowchart (Figure 1).

Figure 1. The consolidated standards of reporting trial chart. MEMS: Medical Event Monitoring System; MIST: Mobile Interactive Supervised Therapy; PC: pill count; SR: 7-day recall self-report.



Mobile Interactive Supervised Therapy Development and Intervention

MIST was developed based on the Theory of Planned Behavior. The MIST system consisted of three components: electronic reminder notifications from the health care system to the patient, transmissions of a daily video of pill-taking from the patient to the health care system, and a color-coded calendar tracker to allow the patient to compare his or her adherence data on any given day with the past week, month, and overall. Screenshots of the MIST features can be found in Figure 2. The previous prototype was developed for both Android and iOS operating systems. Our past pilot study using the earlier prototype showed that MIST was useful in measuring adherence and was acceptable to 42 healthy volunteers [26]. However, as MIST was still a pilot system focusing on producing a functional system, some bugs were identified during pilot testing. Further work was needed to improve the design and usability of the system to ensure the system was robust across a range of users. Since then, MIST has been further developed including fixing bugs such as video upload failure and app instability and was tested among a wide spectrum of phone models during 2016 to 2017. In addition, a reward system was incorporated to incentivize adherence. Before use in this study, the new version was tested in 10 healthy volunteers to ensure technical feasibility.

Participants in the intervention group were asked to use MIST, which comprised integrated Web- and mobile phone–based components. The Web component of the system allowed the administrator to set the time of pill taking each day and to review the submitted videos. The time of the reminder was tagged to the participants' existing daily ART dosing schedules. Once a daily pill dose time was set, the Web system sent pill reminders (via an SMS or a push notification) to the participant's mobile phone 15 min before the pill time, right on the pill time, and half an hour after the pill time until a video was received. If no response was received within half an hour after the pill time, the notifications would stop, and the video would be logged as *missing*. When the participants logged their medications on time (between 15 min before and after pill set timing), the frequency of reminders decreased. The Web system stored and uploaded date- and time-stamped videos onto a secure cloud server linked only to a subject number. Successfully uploaded videos were automatically deleted from the phone. Participants who were randomized to the intervention group received one Xiaomi 4A phone (using the Android system) with the MIST app already installed. The standardized phone allowed the app supporter to rapidly resolve any technical issues experienced by participants.

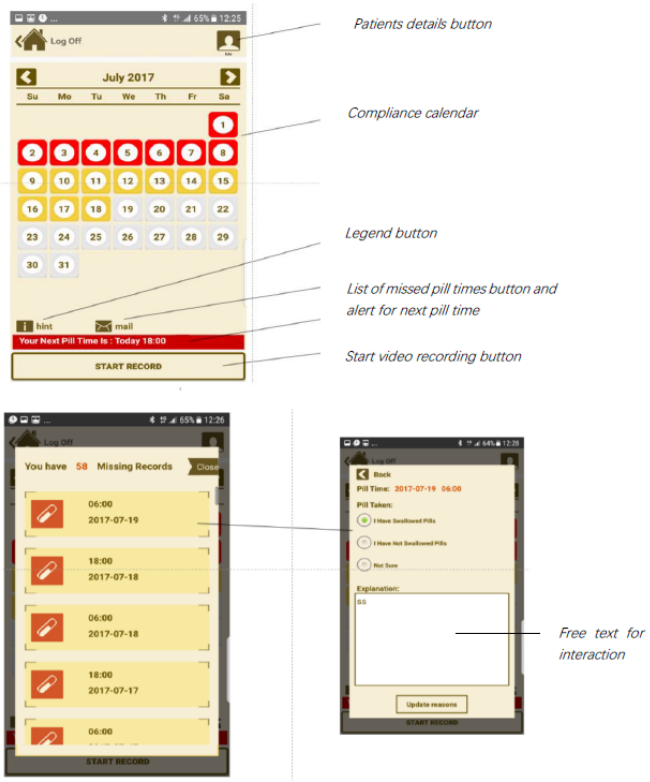
A research assistant (RA) reviewed the uploaded videos daily to ensure that these were of acceptable quality and confirmed pill ingestion. Participants who failed to send a video on time or sent a video that did not permit a verification of pill ingestion were required to provide a reason in their next log-in. If no reason was provided, the RA would contact the participant the next day and ask for an explanation.

The calendar cell colors were synchronized with color codes: a green-colored cell indicated medication adherence, a red-colored cell indicated that the medication was not taken, and a yellow-colored cell indicated that the video was pending the reviewer's grading. The participants could track their compliance by tapping on each cell of the calendar matrix. There

was a built-in reward system that served to reward positive behavior. A reward of Singapore $5 (approximately US $4) was given for every five consecutive successful videos that verified pill intake. The balance could be accumulated and stored in the virtual wallet module of the MIST app and paid to the participants at the final study visit. A maximal reward of Singapore $60 (approximately US $43) would be paid to the MIST users upon completion of the 2-month study.

The participants in the intervention group received a step-by-step guided tour of all the MIST app features via face-to-face demonstrations. All participants verbally indicated that they were confident in using MIST after the demonstrations.

Figure 2. Screenshots of Mobile Interactive Supervised Therapy features.



Outcome Measures and Instruments

The HIV medication adherence rates during the 2-month study period were measured with a combination of objective (pill count [PC], Medication Event Monitoring System [MEMS], and MIST) and subjective (7-day recall self-report [SR]) measures as suggested by the literature [27,28]. We considered taking at least 95% of the prescribed medication as a cut-off to define good adherence in this study as suggested by the literature [29,30].

7-Day Recall Self-Report

Participants were asked about the number of HIV medications they missed in the preceding 7 days at the baseline visit and during the two follow-up visits (1-month visit and 2-month visit). The self-reported medication adherence rate was calculated by subtracting the number of prescribed doses and the number of self-reported missed doses and dividing it by the number of prescribed doses in the past 7 days.

Pill Count

The HIV adherence rate was calculated by subtracting the number of prescribed doses in the past 30 days and the number of missed doses in the past 30 days and dividing it by the number of prescribed doses in the past 30 days [31].

Medication Event Monitor System

MEMSCap Medication Event Monitoring System, developed by Aardex Group, has a pill bottle with a special cap that contains integrated microcircuits to record the date and time whenever a patient opens a bottle, which wirelessly transfers dosing data when used in conjunction with an MEMSCap reader. A secure, Web-based data platform, AARDEX, processes the transferred adherence data using validated algorithms and presents the information in easy-to-interpret graphs and tables [32-34]. The medication adherence rate is calculated based on the number of times the MEMSCap has been opened divided by the number of times the MEMSCap is scheduled to be opened.

Mobile Interactive Supervised Therapy App

Our previous work has demonstrated that MIST is able to accurately measure medication adherence among healthy volunteers in a small pilot study [26]. The ART adherence rate measured by MIST is calculated based on the number of videos received confirming pill intake plus the number of videos not received due to technical problems divided by the number of videos expected during the 2 months.

Process Evaluation

Acceptability and perceptions on the use of the MIST app were assessed during the final study visit. The RA, who was bilingual in English and Chinese and trained in qualitative research, conducted the semistructured interviews. The interview questions were developed based on the research aims in addition to literature review. Two experts were invited to review the interview guide's contents upon its development, and minor revisions were made thereafter. The interview was pilot-tested on 3 healthy volunteers to check the appropriateness of the questions as well as the interview process. Further revisions were made to make the questionnaire understandable. The interview mainly addressed five areas: perceptions of the overall MIST app user experience, perceived usefulness of each feature, strengths and weaknesses of MIST, interest in continuing to use MIST in the future, and recommendations for further improvements. A total of 20 individual face-to-face interviews were conducted, lasting for 3 to 13 min.

Study Procedure

Data collection commenced after ethical approval had been obtained. The primary physicians of the potential recruits were responsible for introducing the participants to the RA. The RA passed a short list of potential recruits to their primary physicians during their scheduled clinic appointments. The primary physicians would briefly introduce the study to their patients. Those who showed interest were then introduced to the RA who would further explain the study to them in a separate private room. Only after obtaining a written consent would the potential participant be recruited into the study. This was followed by a collection of demographic and clinical data and baseline data via a self-administered questionnaire before randomization took place. Outcomes were measured at the following time points for all participants from the two groups: (1) 1 month after the intervention (posttest 1) and (2) 2 months after the intervention (posttest 2).

The RA set up a unique account that was password protected for each participant in the intervention group to log into the MIST app at the baseline visit. The MIST app was activated in the provided Xiaomi 4A mobile phones, and the participants were provided with detailed face-to-face instructions on how to use it. The same RA was responsible for the collection of all data, including process evaluation interviews.

Data Analysis

Quantitative analyses were performed using the Statistical Package for the Social Sciences (IBM, Version 21.0 [35]).$P<.05$ was considered statistically significant. Descriptive statistics were used to describe demographic data, clinical data, and baseline self-reported medication adherence rates for both groups. Either the chi-square test, Fisher exact test, or the Mann-Whitney *U* test was used to assure comparability between the intervention and control groups.

To answer hypothesis 1, the Mann-Whitney *U* test was used to examine differences in ART adherence rates as measured by the 7-day recall SR, PC, and MEMS between the two groups at the 1-month and 2-month visits.

To answer hypothesis 2, the Wilcoxon signed-rank test was used to compare ART adherence rates (measured by PC, MEMS, and with or without MIST) at different time points (1-month and 2-month visits) for each group, and the Friedman test was used to compare the ART adherence rates (measured by SR) at baseline and the two follow-up visits.

Qualitative data from the interviews were analyzed using content analysis. The audio-taped interview data were transcribed verbatim by the RA to capture nonverbal nuances. A range of highlighter colors were used to highlight segments of data portraying similar ideas and meaning. Subsequently, data extracts that have been coded with the same color were collated and organized into categories. The refined potentially repeated patterns across the entire data were, then, transferred to a separated word document. The initial codes were reviewed again by the RA and another investigator, and related codes were collated to form subthemes [36]. Different opinions were discussed until a mutual agreement was reached between the two investigators [37]. Rigor was ensured through credibility, transferability, dependability, and confirmability in this study process [38].

Ethical Considerations

Ethics approval for the study was received from the institutional review board before commencing the study (Ref: 2017/00150). All participants provided written consent for audio-recordings of the interviews. Confidentiality of the data and voluntary participation were explained to all participants.

Results

Participants' Demographics and Clinical Data and Group Comparison

Among the 40 participants who participated in the study, 38 (95%) were male and 25 (63%) were Chinese, 8 (20%) were Malay, 3 (8%) were Indian, and 4 (10%) were from other ethnic groups. The median age was 37.45 (IQR 30.07-44.95) years. In total, 21 of 40 (53%) participants were doing professional work. Nearly half of them (19/40, 48%) completed university or higher education. Most were single (32/40, 80%), lived with family (29/40, 73%), lived in a Housing and Development Board flat or a studio apartment (36/40, 90%), and had home Wi-Fi (29/40, 73%). The median ART durations at the baseline were 25.15 (IQR 12.23-48.03) months and 22.35 (IQR 8.93-43.28) months for the control and intervention groups, respectively. Taking two types of tablets of HIV medication was the most commonly reported regime (16/40, 40%) among the participants, followed by one type of tablet (11/40, 28%), three types of tablets (8/40, 20%), and at least four types of tablets (5/40, 13%). The vast majority of the participants were very confident in operating smartphones (34/40, 85%). As shown in Multimedia Appendix 1, baseline characteristics were similar between the control and interventional groups except for types of ART medication ($P=.01$).

Comparison of Outcomes

As shown in Table 1, there were no significant differences in the medication adherence rates between the intervention and control groups at the follow-up time points. In addition, neither

of the measures reported significant differences in ART adherence rates between the two groups across the 2-month period (Table 2).

Table 1. Comparison of median antiretroviral therapy adherence rates between the control group and intervention group at the 1-month and 2-month follow-up visits (N=40).

Follow-up visit and ART[a] adherence rate (%) measurement	Control group, median (IQR)	Intervention group, median (IQR)	*P* value
1 month			
Self-report	100 (100-100)	100 (100-100)	>.99
Pill count	100 (100-100)	100 (100-100)	.86
MEMS[b] median	100 (93-100)	100 (93-100)	.74
2 months			
Self-report	100 (100-100)	100 (100-100)	.80
Pill count	100 (100-100)	100 (100-100)	.84
MEMS median	97 (93-100)	97.5 (92-100)	.82

[a]ART: antiretroviral therapy.

[b]MEMS: Medication Event Monitoring System.

Table 2. Comparison of median antiretroviral therapy adherence rates at different time points for each group (N=40).

ART[a] adherence rate (%) measurement	Baseline, median (IQR)	1-month visit, median (IQR)	2-month visit, median (IQR)	*P* value
Self-report				.22[b]
Control group	100 (100-100)	100 (100-100)	100 (100-100)	
MIST[c] intervention group	100 (100-100)	100 (100-100)	100 (100-100)	
Pill count				.07[d]
Control group	N/A[e]	100 (100-100)	100 (100-100)	
MIST intervention group	N/A	100 (100-100)	100 (100-100)	
MEMS[f]				.63[d]
Control group	N/A	100 (93-100)	97 (93-100)	
MIST intervention group	N/A	100 (93-100)	97.5 (92-100)	
MIST				
MIST intervention group	N/A	96 (87-100)	94 (87-99)	.25[d]

[a]ART: antiretroviral therapy.

[b]Friedman test.

[c]MIST: Mobile Interactive Supervised Therapy.

[d]Wilcoxon signed-rank test.

[e]N/A: Not applicable.

[f]MEMS: Medication Event Monitoring System.

Process Evaluation

All 20 participants from the intervention group completed the semistructured interviews. Among 20 participants who were interviewed, the majority of them were extremely or very confident (18/20, 90%) in operating a mobile phone. Participants were aged between 27 and 52 years. Most were males (n=19) and Chinese (n=13). Half of them were doing professional work, 2 were doing skilled work, while 7 were unemployed, and 1 was a student. Slightly over half of them completed university or higher education, 35% (7/20) completed a college, and the remaining 10% (2/20) received a secondary school education. The majority of them were single (16/20, 80%) and lived with family or friends. Findings related to the MIST app user experiences were grouped into six categories: (1) perceived ease of use, (2) benefits of MIST, (3) MIST app features preferences, (4) MIST app dislikes, (5) future willingness to use MIST, and (6) suggestions for MIST improvement.

Category 1: Perceived Ease of Use

When asked about their overall experience with the MIST app, half of the participants reported that the app was user-friendly.

Category 2: Benefits of Mobile Interactive Supervised Therapy

Most of the participants considered MIST to be beneficial. It not only reminded and motivated them to take their ART medications punctually and consistently every day but also helped in situations when they were tired and forgot to take their medications. They also agreed the color-coded calendar helped them to track their medication-taking histories. In all, 2 participants verbalized that receiving cash helped to lighten their financial burdens.

Category 3: Mobile Interactive Supervised Therapy App Feature Preferences

When asked for their preferred app features, the participants voted almost in equal numbers for each feature, that is, SMS reminders, color-coded calendars, and the reward system were voted by 7, 6, and 5 participants, respectively.

Categories 4 and 5: Mobile Interactive Supervised Therapy App Dislikes and Willingness for Future Utilization

Some participants revealed that there were certain aspects of the MIST app that they disliked, and these included the need for an internet connection, app glitches and instability, limitation to mobile phones using the Android platform, extra burdens (such as recording themselves taking their medications), and privacy concerns. Due to these concerns, among the 20 participants who were interviewed, only 6 (30%) participants expressed their willingness to use MIST in the future.

As there was stigma associated HIV infections, the participants tried to minimize receiving SMS reminders or taking their medications in public to prevent attracting attention. For instance, 1 participant switched off the provided phone as he was unsure what the contents of the incoming SMS reminders would look like. Another participant was concerned that recording himself taking the medication would trigger curious onlookers. A third participant suggested that MIST should have an option that allowed the users to change the app icon image as the default icon might inadvertently reveal their HIV status.

Category 6: Suggestion for Mobile Interactive Supervised Therapy Improvement

Most of the participants suggested making the app features more user-friendly. The MIST app should be made available for all types of phones. They also suggested alternative reward methods and methods to track the medications taken. Some thought that the app should incorporate a fun element to engage future users.

Discussion

Principal Findings

The pilot trial on the use of MIST, which followed patients infected with HIV for 2 months, did not show any significant improvements in ART adherence rates. However, this finding needs to be interpreted with caution as this was just a pilot study and the sample recruited were those who had good medication adherence. On the basis of the process evaluation, MIST was perceived to be beneficial among our participants. Our process evaluation provided useful data to further develop MIST for bigger and long-term mobile phone app–assisted intervention RCTs in the future. The reason for the lack of improvement was the ceiling effect. Before the study, both groups of participants were already adhering to their medications. For example, the median 7-day self-recalled ART adherence at the baseline was 100% in both groups, leaving no room for them to improve their adherence further. Given a longer enrollment time, we would have liked to target subjects who were at high risk of nonadherence. However, our subject recruitment was limited by the small number of newly diagnosed patients annually. In addition, the recruitment was hampered by perceived stigma in people living with HIV and concern that the intervention might compromise confidentiality and privacy. After a slow start of having only 3 patients recruited, we expanded our eligibility criteria to include those who had no adherence problems as well. However, the rejection rate was still nearly 48%. The study ended with most of the participants already having good ART adherence. The selection bias was likely to contribute to our observation that patients with poor ART adherence might be also those who were less interested to participate in the trial. In support of our assertion of the ceiling effect, one study [15] did not have the ceiling effect as it only enrolled participants who reported less than 95% adherence and demonstrated adherence improvements following their mobile phone app use. The second possible reason for the lack of improvement in our study is that most of our participants were already on ART for about 1 year, as evidenced by the median ART duration (IQR) for the control group of 25 months (IQR 12-48 months) and the intervention group of 22 months (IQR 8-43 months) before the study enrollment. The need for lifelong daily ART adherence would have developed among them a well-established daily medication-taking routine [39]. In all, two previous phone technology–assisted adherence RCTs [40,41], which targeted only patients newly initiated onto ART, demonstrated the effectiveness of SMS or SMS combined with counseling in improving medication adherence. The results highlighted the importance of initiating the intervention at the early stage before the participants establish medication-taking habits [40,41].

Despite the absence of an effect of MIST on ART adherence among the participants who had 100% adherence rates at baseline, our qualitative findings demonstrated that the MIST app was beneficial to most participants. They perceived MIST to be user-friendly. It is interesting to note that although most of our participants did not forget to take their medications, the reminders were especially helpful for others who did shift work or who were busy with other things at the time their medications were due. This finding is consistent with a previous RCT [39], in which some participants who did not benefit from reminders would nevertheless recommend it to someone who already had a fixed HIV medication–taking routine [39]. Some perceived the color-coded calendar tracking system to be more useful compared with a conventional tracking tool, such as a diary or pillbox, or even no tracking tool. The visual calendar medication-taking tracking feature helped our participants to save time and, more importantly, motivated them to continue their good medication-taking behaviors. Our study findings are consistent with previous research, which suggested that an app

with health-related behaviors and goal tracking features, such as medication-taking behaviors, was perceived as valuable [42].

We believe the following reasons contributed to the reluctance of most of the participants (14/20) in continuing to use MIST in the future. First, the availability of the latest prototype version was limited to the Android operating system as it was still being optimized across all Android operating system phones during the study period. Owing to this limitation, we had to provide each participant with a Xiaomi 4A model phone that contained the MIST app. Failure to carry the phone, turn it on, or have the battery charged prevented the reminders from serving their desired purpose. Second, occasional app glitches and instability were barriers that discouraged participants from continuing to use the MIST app. Some participants claimed that they would rather have taken their medications than wait for the app to start working. Third, some participants found video recording while taking their medications burdensome, which might not be sustainable as HIV requires lifelong medication adherence. However, in spite of this, all 20 participants maintained their commitments and continued using MIST until the end of the study. On the contrary, a trial [13] observed a 25% reduction in the use of their Heart2HAART mobile phone app during a 3-month trial period. In their study, participants in the Heart2HAART group were expected to respond to randomly generated daily medication prompts, including questions about HIV medication side effects and substance cravings during the study period. They concluded that user fatigue might have contributed to app usage reduction [13]. Our study finding stood in contrast to the previous study findings [13], possibly because of the cash reward system that was built into MIST, whereas there was no incentive system in the previous study [13]. As having cash incentives might not be a practical measure in the long term, future studies should consider alternative incentive methods.

Although the MIST app is password protected, some of the participants expressed concerns about possible disclosures of their HIV status during the interviews. Compounding the privacy concern was the bulky electronic monitor device—MEMS—which might have triggered onlookers' curiosity and a loss of privacy thereafter. Our findings are consistent with findings from other studies. For example, a global qualitative review study revealed that the acceptability of mobile technology–assisted interventions that aimed to improve ART adherence among people living with HIV could be affected by privacy and confidentiality issues [43]. In view of the findings of this study, we will optimize MIST to make sure that it is more secure. We will modify the MIST prototype based on other suggestions of our participants. For instance, the app icon should be made customizable. MIST should be optimized to market-level quality, meaning that it should be made more sensitive and less glitchy and it should be downloadable for use in other operating systems. We will also develop a backup plan in case the MIST app fails to work. For example, verbal reports or SMS as an alternative to report adherence.

Limitations of the Study

This study had several limitations. First, owing to pragmatic considerations, the small sample size of 40 participants might have failed to detect a significant difference in medication adherence (ie, type 2 error). Second, we did not collect biological data such as CD4 count and viral load, which are good indicators for treatment success. Third, although this study was conducted in a tertiary hospital, due to the intervention itself, some older people were excluded from this study as they were not confident in using the technology. Blinding was not possible for both participants and researchers due to the nature of the app. Hence, this study might have introduced some response bias. MIST participants might have responded differently owing to social expectation pressure. Fourth, the interview duration was short. However, this was a process evaluation as opposed to an in-depth qualitative study, and all interview questions were straightforward. Furthermore, the majority of potential participants declined to be interviewed owing to stigma. Stigma could also have limited the information that enrolled participants were willing to share during interviews. Finally, because of time constraints, the majority of our participants had good baseline ART adherence resulting in the ceiling effect. The intervention was limited to a 2-month follow-up, and this short follow-up period does not reflect the real-world situation, as HIV requires lifelong commitment and adherence.

Conclusions

In conclusion, this pilot study showed that MIST was perceived as beneficial by our participants. Future RCTs with a bigger sample size and the use of an optimized and more acceptable version of MIST should be conducted to evaluate the effectiveness of MIST in enhancing medication adherence among people living with HIV, especially those who have poor adherence and those who have just started on ART.

Acknowledgments

The authors appreciate the National Health Innovation Centre (NHIC) and the National University Hospital Singapore who provided funding for this study. The authors thank all participants and clinic staff members for their support and contributions. They also thank the National University Health System Medical Publications Support Unit for assistance in language editing of this manuscript. This study was supported by the NHIC Innovation to Develop Grant (grant number: NHIC-I2D-1509082) and the National University Health System Allied Health and Nursing Grant (grant number: NUHRO/2018/043/AHN/01).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparison of the participants' sociodemographic, clinical, and technical skills data (N=40).
[DOCX File , 20 KB - mhealth_v8i3e15702_app1.docx]

Multimedia Appendix 2
CONSORT-EHEALTH checklist V 1.6.1.
[PDF File (Adobe PDF File), 2344 KB - mhealth_v8i3e15702_app2.pdf]

References

1. UNAIDS. Joint United Nations Programme on HIV/AIDS. 2017. UNAIDS Data 2018 URL: https://www.unaids.org/sites/default/files/media_asset/unaids-data-2018_en.pdf
2. Ministry of Health Singapore. Update on the HIV/AIDS situation in Singapore 2017 (JUNE 2018) URL: https://www.moh.gov.sg/resources-statistics/infectious-disease-statistics/hiv-stats/update-on-the-hiv-aids-situation-in-singapore-2017-(june-2018) [accessed 2019-01-20]
3. Mills EJ, Lester R, Thorlund K, Lorenzi M, Muldoon K, Kanters S, et al. Interventions to promote adherence to antiretroviral therapy in Africa: a network meta-analysis. Lancet HIV 2014 Dec;1(3):e104-e111 [FREE Full text] [doi: 10.1016/S2352-3018(14)00003-4] [Medline: 26424119]
4. Braitstein P, Brinkhof MW, Dabis F, Schechter M, Boulle A, Miotti P, Antiretroviral Therapy in Lower Income Countries (ART-LINC) Collaboration, ART Cohort Collaboration (ART-CC) groups. Mortality of HIV-1-infected patients in the first year of antiretroviral therapy: comparison between low-income and high-income countries. Lancet 2006 Mar 11;367(9513):817-824. [doi: 10.1016/S0140-6736(06)68337-2] [Medline: 16530575]
5. Gill CJ, Hamer DH, Simon JL, Thea DM, Sabin LL. No room for complacency about adherence to antiretroviral therapy in sub-Saharan Africa. AIDS 2005 Aug 12;19(12):1243-1249 [FREE Full text] [doi: 10.1097/01.aids.0000180094.04652.3b] [Medline: 16052078]
6. de Olalla Garcia P, Knobel H, Carmona A, Guelar A, López-Colomés JL, Caylà JA. Impact of adherence and highly active antiretroviral therapy on survival in HIV-infected patients. J Acquir Immune Defic Syndr 2002 May 1;30(1):105-110. [doi: 10.1097/00042560-200205010-00014] [Medline: 12048370]
7. Bangsberg DR, Deeks SG. Is average adherence to HIV antiretroviral therapy enough? J Gen Intern Med 2002 Oct;17(10):812-813 [FREE Full text] [doi: 10.1046/j.1525-1497.2002.20812.x] [Medline: 12390559]
8. Bartlett JA. Addressing the challenges of adherence. J Acquir Immune Defic Syndr 2002 Feb 1;29(Suppl 1):S2-10. [doi: 10.1097/00126334-200202011-00002] [Medline: 11832696]
9. Ortego C, Huedo-Medina TB, Llorca J, Sevilla L, Santos P, Rodríguez E, et al. Adherence to highly active antiretroviral therapy (HAART): a meta-analysis. AIDS Behav 2011 Oct;15(7):1381-1396. [doi: 10.1007/s10461-011-9942-x] [Medline: 21468660]
10. Haase J, Farris KB, Dorsch MP. Mobile applications to improve medication adherence. Telemed J E Health 2017 Feb;23(2):75-79. [doi: 10.1089/tmj.2015.0227] [Medline: 27248315]
11. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile phone apps to improve medication adherence: a systematic stepwise process to identify high-quality apps. JMIR Mhealth Uhealth 2016 Dec 2;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
12. Perera AI, Thomas MG, Moore JO, Faasse K, Petrie KJ. Effect of a smartphone application incorporating personalized health-related imagery on adherence to antiretroviral therapy: a randomized clinical trial. AIDS Patient Care STDS 2014 Nov;28(11):579-586 [FREE Full text] [doi: 10.1089/apc.2014.0156] [Medline: 25290556]
13. Himelhoch S, Kreyenbuhl J, Palmer-Bacon J, Chu M, Brown C, Potts W. Pilot feasibility study of Heart2HAART: a smartphone application to assist with adherence among substance users living with HIV. AIDS Care 2017 Jul;29(7):898-904. [doi: 10.1080/09540121.2016.1259454] [Medline: 28043176]
14. Hightow-Weidman L, Muessig K, Knudtson K, Srivatsa M, Lawrence E, LeGrand S, et al. A gamified smartphone app to support engagement in care and medication adherence for HIV-positive young men who have sex with men (AllyQuest): development and pilot study. JMIR Public Health Surveill 2018 Apr 30;4(2):e34 [FREE Full text] [doi: 10.2196/publichealth.8923] [Medline: 29712626]
15. Martin CA, Upvall MJ. A mobile phone HIV medication adherence intervention: acceptability and feasibility study. J Assoc Nurses AIDS Care 2016;27(6):804-816. [doi: 10.1016/j.jana.2016.07.002] [Medline: 27497514]
16. Dworkin M, Chakraborty A, Lee S, Monahan C, Hightow-Weidman L, Garofalo R, et al. A realistic talking human embodied agent mobile phone intervention to promote HIV medication adherence and retention in care in young HIV-positive African American men who have sex with men: qualitative study. JMIR Mhealth Uhealth 2018 Jul 31;6(7):e10211 [FREE Full text] [doi: 10.2196/10211] [Medline: 30064971]
17. Westergaard RP, Genz A, Panico K, Surkan PJ, Keruly J, Hutton HE, et al. Acceptability of a mobile health intervention to enhance HIV care coordination for patients with substance use disorders. Addict Sci Clin Pract 2017 Apr 26;12(1):11 [FREE Full text] [doi: 10.1186/s13722-017-0076-y] [Medline: 28441962]

18. Julious S, Machin D, Tan SB. An Introduction to Statistics in Early Phase Trials. Hoboken, New Jersey, United States: Wiley; 2010.
19. Billingham SA, Whitehead AL, Julious SA. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database. BMC Med Res Methodol 2013 Aug 20;13:104 [FREE Full text] [doi: 10.1186/1471-2288-13-104] [Medline: 23961782]
20. Kieser M, Wassmer G. On the use of the upper confidence limit for the variance from a pilot sample for sample size determination. Biomed J 1996;38(8):941-949. [doi: 10.1002/bimj.4710380806]
21. Matthews JN. Introduction to Randomized Controlled Clinical Trials. Milton Park, Abingdon, Oxfordshire: Routledge; 2006.
22. Research Randomizer. URL: https://www.randomizer.org/ [accessed 2018-03-15]
23. Doig GS, Simpson F. Randomization and allocation concealment: a practical guide for researchers. J Crit Care 2005 Jun;20(2):187-91; discussion 191. [doi: 10.1016/j.jcrc.2005.04.005] [Medline: 16139163]
24. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet 2002 Feb 16;359(9306):614-618. [doi: 10.1016/S0140-6736(02)07750-4] [Medline: 11867132]
25. 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]
26. Molton JS, Pang Y, Wang ZC, Qiu BQ, Wu P, Rahman-Shepherd A, et al. Prospective single-arm interventional pilot study to assess a smartphone-based system for measuring and supporting adherence to medication. BMJ Open 2016 Dec 20;6(12):e014194 [FREE Full text] [doi: 10.1136/bmjopen-2016-014194] [Medline: 27998903]
27. Farmer KC. Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice. Clin Ther 1999 Jun;21(6):1074-90; discussion 1073. [doi: 10.1016/S0149-2918(99)80026-5] [Medline: 10440628]
28. Lam WY, Fresco P. Medication adherence measures: an overview. Biomed Res Int 2015;2015:217047 [FREE Full text] [doi: 10.1155/2015/217047] [Medline: 26539470]
29. Paterson DL, Swindells S, Mohr J, Brester M, Vergis EN, Squier C, et al. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann Intern Med 2000 Jul 4;133(1):21-30. [doi: 10.7326/0003-4819-133-1-200007040-00004] [Medline: 10877736]
30. Bangsberg DR, Hecht FM, Charlebois ED, Zolopa AR, Holodniy M, Sheiner L, et al. Adherence to protease inhibitors, HIV-1 viral load, and development of drug resistance in an indigent population. AIDS 2000 Mar 10;14(4):357-366. [doi: 10.1097/00002030-200003100-00008] [Medline: 10770537]
31. Castillo-Mancilla JR, Haberer JE. Adherence measurements in HIV: new advancements in pharmacologic methods and real-time monitoring. Curr HIV/AIDS Rep 2018 Feb;15(1):49-59 [FREE Full text] [doi: 10.1007/s11904-018-0377-0] [Medline: 29380227]
32. Shuter J, Sarlo JA, Kanmaz TJ, Rode RA, Zingman BS. HIV-infected patients receiving lopinavir/ritonavir-based antiretroviral therapy achieve high rates of virologic suppression despite adherence rates less than 95%. J Acquir Immune Defic Syndr 2007 May 1;45(1):4-8. [doi: 10.1097/QAI.0b013e318050d8c2] [Medline: 17460469]
33. Gross R, Bellamy SL, Chapman J, Han X, O'Duor J, Palmer SC, et al. Managed problem solving for antiretroviral therapy adherence: a randomized trial. JAMA Intern Med 2013 Feb 25;173(4):300-306 [FREE Full text] [doi: 10.1001/jamainternmed.2013.2152] [Medline: 23358784]
34. Hardy H, Kumar V, Doros G, Farmer E, Drainoni M, Rybin D, et al. Randomized controlled trial of a personalized cellular phone reminder system to enhance adherence to antiretroviral therapy. AIDS Patient Care STDS 2011 Mar;25(3):153-161 [FREE Full text] [doi: 10.1089/apc.2010.0006] [Medline: 21323532]
35. IBM Corp. IBM. Armonk, NY: IBM Corp; 2012. IBM SPSS Statistics 21 Documentation URL: https://www.ibm.com/support/pages/ibm-spss-statistics-21-documentation [accessed 2019-02-26]
36. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
37. Krefting L. Rigor in qualitative research: the assessment of trustworthiness. Am J Occup Ther 1991 Mar;45(3):214-222. [doi: 10.5014/ajot.45.3.214] [Medline: 2031523]
38. Guba EG, Lincoln YS. Competing paradigms in qualitative research. In: Handbook of Qualitative Research. Thousand Oaks, California, United States: Sage Publications; 1994:105-117.
39. Mao LM, Buchanan A, Wong HTH, Persson A. Beyond mere pill taking: SMS reminders for HIV treatment adherence delivered to mobile phones of clients in a community support network in Australia. Health Soc Care Community 2018 Jul;26(4):486-494. [doi: 10.1111/hsc.12544] [Medline: 29336111]
40. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]
41. Abdulrahman SA, Rampal L, Ibrahim F, Radhakrishnan AP, Shahar HK, Othman N. Mobile phone reminders and peer counseling improve adherence and treatment outcomes of patients on ART in Malaysia: A randomized clinical trial. PLoS One 2017;12(5):e0177698 [FREE Full text] [doi: 10.1371/journal.pone.0177698] [Medline: 28520768]

42. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
43. Ma QY, Tso LS, Rich ZC, Hall BJ, Beanland R, Li H, et al. Barriers and facilitators of interventions for improving antiretroviral therapy adherence: a systematic review of global qualitative evidence. J Int AIDS Soc 2016;19(1):21166 [FREE Full text] [doi: 10.7448/IAS.19.1.21166] [Medline: 27756450]

Abbreviations

ART: antiretroviral therapy
MEMS: Medication Event Monitoring System
MIST: Mobile Interactive Supervised Therapy
NHIC: National Health Innovation Centre
PC: pill count
RA: research assistant
RCT: randomized controlled trial
SR: self-report

Edited by G Eysenbach; submitted 31.07.19; peer-reviewed by J Zhu, C Sun; comments to author 04.10.19; revised version received 29.11.19; accepted 16.12.19; published 27.03.20.

Please cite as:
Pang Y, Molton JS, Ooi WT, Paton NI, He HG
Preliminary Effects of a Mobile Interactive Supervised Therapy Intervention on People Living With HIV: Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(3):e15702
URL: http://mhealth.jmir.org/2020/3/e15702/
doi:10.2196/15702
PMID:32217500

Original Paper

Importance of Photography Education to Improve Image Quality for Accurate Remote Diagnoses in Dental Trauma Patients: Observational Study

Jin-Sun Jeong[1,2*], BA, PhD; Nan-Sim Pang[3*], MSD, DDS, PhD; Yiseul Choi[3,4], BS; Kyeong-Mee Park[3,4], BS, PhD; Taekbin Kim[3], DDS; Xin Xu[1,2], MSD, DDS, PhD; Wonse Park[3], MSD, DDS, PhD

[1]School and Hospital of Stomatology, Shandong University, Jinan, China

[2]Shandong Key Laboratory of Oral Tissue Regeneration & Shandong Engineering Laboratory for Dental Materials and Oral Tissue Regeneration, Jinan, China

[3]Department of Advanced General Dentistry, College of Dentistry, Yonsei University, Seoul, Republic of Korea

[4]Human Identification Research Institute, College of Dentistry, Yonsei University, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Wonse Park, MSD, DDS, PhD
Department of Advanced General Dentistry
College of Dentistry
Yonsei University
50-1, Yonsei-ro
Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2228 8980
Email: wonse@yuhs.ac

Abstract

Background: High-quality photos are critical for the remote diagnosis of dental trauma and thus are beneficial to the prognosis. The quality of the images obtained using a cell phone depends on the level of dental and photography knowledge of the person who is taking the photos.

Objective: This study aimed to determine the efficacy of photography education in improving images used for the remote diagnosis of dental trauma.

Methods: The subjects comprised 30 laypeople and 30 dentists who were randomly assigned to 15 subgroups with 2 subjects in each. Each subject was asked to take photos of their own anterior teeth and those of their partner on the assumption that an accident occurred using both an iPhone 4s and iPhone 6. Education about how to take an appropriate photo of the anterior teeth for teleconsultation purposes was then provided, after which photos were taken again. Photos were assessed by a dentist for their usefulness in diagnosis.

Results: This study analyzed 965 photos: 441 taken by laypeople and 524 taken by dentists. Photos taken after providing education had significantly higher scores for all assessment items than those taken before education (P<.05). The scores were also significantly higher for photos taken using the rear camera than those taken using the front camera (P<.02). The iPhone 6 did not have overwhelming advantages. The photos taken by dentists had significantly higher scores than those taken by laypeople for most of the evaluated items.

Conclusions: Both laypeople and dentists might find photography education useful for when they are taking photos to be used in teleconsultations. The type of cell phone does not significantly affect the usefulness of such photos.

(JMIR Mhealth Uhealth 2020;8(3):e15152) doi:10.2196/15152

KEYWORDS

telemedicine; remote consultation; emergencies; tooth injuries; cell phone

Introduction

The World Health Organization has recently released a statement emphasizing the use of appropriate digital technologies for public health [1]. Ubiquitous health care (uHealth) services are rapidly developing due to increasing attention worldwide [2]. UHealth refers to combining information technology with medical services in order to provide remote medical and health management services that can be used anytime and anywhere [3]. UHealth allows for remote medical services such as remote examinations of and prescribing for various medical conditions as well as remote health management and enhancement services for healthy clients [4].

Rapid progress in information technology has resulted in successful implementation and testing of the electronic submission of clinical images for remote consultations in most medical and surgical subspecialties [5]. Moreover, implementations and research have been carried out using patient-to-doctor remote consultations, called telemedicine [6]. Home telenursing for patients suffering heart failure or diabetes, teleradiology for ultrasound and x-ray images, and teleconsultation for emergency orthopedic patients have been widely studied [7-12]. Studies related to dentistry have investigated remote medical procedures such as orthodontic consultations, remote oral care, preoperative evaluations before implant placement, and treatment of traumatic tooth injuries using teledentistry [13-16]. In cases of trauma to teeth or alveolar bone, the prognosis is most significantly affected by how rapidly a diagnosis is made and treatment is applied [17,18]. These considerations warrant the development of remote medical services allowing rapid diagnosis and treatment in the event of traumatic tooth injuries occurring when a dentist or other tooth injury expert is not available nearby.

Devices used to provide remote oral care services include oral video cameras (intraoral image capturing devices), digital single lens reflex (DSLR) cameras, and cell phones with built-in cameras [19-21]. An oral video camera is a small device that is easy to use but has the limitations of high-quality images being difficult to obtain due to its size and requiring a computer to display the images. In contrast, a DSLR camera allows for acquisition of images of exceptional quality, but it is relatively expensive, requires special imaging equipment, and the operator needs to be trained; hence, it is not an optimal imaging device for emergency or remote consultations.

Cell phones are highly portable, have high-resolution cameras, and are now almost ubiquitously used by people of all age groups. Continuing developments have allowed for such devices to be used not only for telecommunication but also to provide multiple computer-like communication functions including text, photo, and video transfer, as well as internet access. Different types of remote consultation or treatment based on the use of cell phones are currently being investigated in the field of telemedicine [22,23]. For teleconsultations, cell phone cameras are better than DSLR and oral video cameras in terms of convenience and portability [24].

However, the quality of the images obtained using a cell phone depends on the level of dental knowledge of the person who is taking the photos and the imaging conditions [16]. Zaror et al [25] studied the efficacy of an app for traumatic dental injuries and the quality of images for clinical purposes, but photography education was not involved. An in vitro study identified several important camera-related factors that could influence the image quality for teledentistry in dentoalveolar trauma, including autofocusing and antimovement functions, and the automatic white-balance function was helpful for detecting the color area [26].

This study investigated differences in image quality according to who photographed the traumatized tooth (dentist vs laypeople), the effect of receiving photography education (before vs after), and what model cell phone was used (including different numbers of camera pixels).

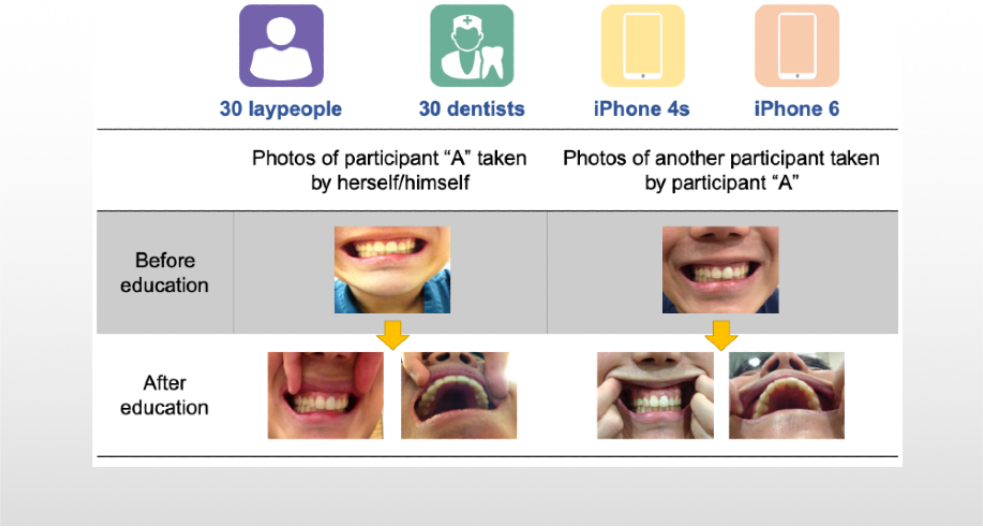
Methods

Recruitment

This clinical research was approved by the institutional review board of the dental hospital at Yonsei University (number 2-2012-0025). The subjects consisted of 60 Korean adults (30 laypeople and 30 dentists) selected from 62 volunteers. The laypeople comprised 21 males and 9 females ranging in age from 21 to 39 years (28.67 [SD 4.20] years), while the dentists comprised 14 males and 16 females who ranged in age from 26 to 37 years (30.74 [SD 3.21] years). The inclusion criteria were being older than 20 years, able to operate the camera on a cell phone, having continuous anterior dentition (including any prosthesis and orthodontic brackets), and signing the consent form for the experiment.

The subjects were randomly arranged into 15 subgroups with 2 subjects each, and they were asked to take photos of themselves and their partner using both an iPhone 4s (0.3-megapixel front camera, 8-megapixel rear camera, Apple), and iPhone 6 (1.2-megapixel front camera, 8-megapixel rear camera, Apple) before and after receiving photography education (Figure 1). There was no limit to the number of photos each subject could take.

Figure 1. Summary of experimental design.



Experimental Design

We assumed that injuries had occurred in the subjects because of an accident. Subjects were encouraged to freely take photos of the maxillary anterior area including 4 incisors and 2 canines by using the functions of autofocusing, blurring, white balance, focal length, and resolution provided by the cell phone, with the built-in front camera and the default camera app. They were then asked to take photos of their colleagues' teeth using the same methods but with the built-in rear camera.

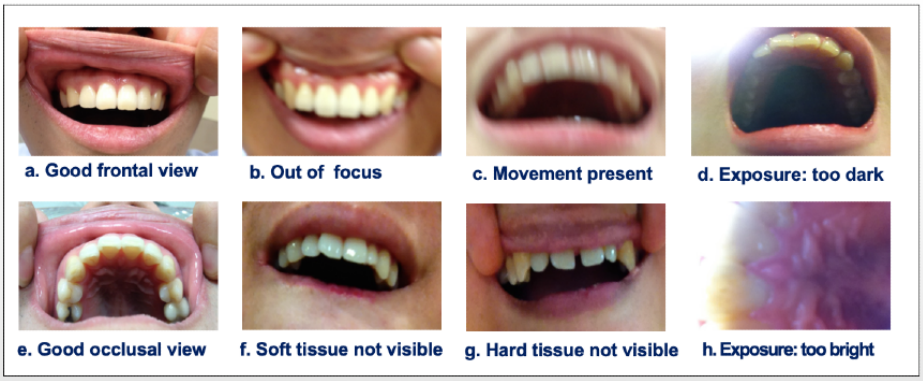
Photography education was then provided by an experienced dental hygienist, face to face, to every subject with the following contents: how to take photos from frontal and occlusal views, how to take a photo in which all 6 anterior teeth appear, how to use a retractor to expose soft tissue, how to adjust the focus, and how to adjust the camera settings according to the protocol proposed by Park et al [26] to better exhibit the anatomy and discoloration of the teeth as well as the gingival texture. The education session lasted about 15 minutes, and the subjects were then asked to repeat the experiment following the instructions provided in the photography education.

Photo Assessments

The captured images were assessed twice with a 1-week interval on a desktop computer (LG, screen resolution 1920×1080 pixels, 8-bit color depth) by a single dentist in the Department of Advanced General Dentistry at Yonsei University.

After all the photos had been collected, errors that appeared frequently during the experiment were analyzed and images of frontal and occlusal views were evaluated. Frequent errors included retraction failure and incomplete frontal and occlusal views. Each photo was evaluated as either good or failure, and the relationships between this categorization and the frequency of errors were evaluated by the dentist (Figure 2). The photos evaluated as good were scored on the following 5-point numerical scale: 1=not suitable for making a diagnosis, 2=able to make a questionable diagnosis, 3=able to make an average diagnosis, 4=able to make a good diagnosis, and 5=able to make a perfect diagnosis.

Figure 2. Examples of photos that were appropriate and inappropriate for the initial assessment.



The two assessments meant that the assessment score for each tooth type was on a 10-point scale. The highest score was used when multiple photos of the same type were assessed. The central incisors, lateral incisors, and canines were assessed separately against the items listed in Table 1 for soft tissue and hard tissue to produce a final assessment score of 30 points.

Table 1. Assessment items.

Category and finding	Suggested situation
General	
Optimal focus	—
Movement present	—
Exposure	—
Hard tissue	
Frontal	
Shape	Crown fracture
Position	Displacement
Alignment	Displacement
Bleeding spot with pink color	Pulp exposure
Occlusal	
Shape	Crown fracture
Position	Displacement
Alignment	Displacement
Bleeding spot with pink color	Pulp exposure
Soft tissue	
Frontal	
Gingival sulcus	Gingival bleeding
Integrity	Laceration
Color	Hematoma
Occlusal	
Gingival sulcus	Gingival bleeding
Integrity	Laceration
Color	Hematoma

Statistical Analysis

Data were analyzed using SPSS Statistics version 25.0 (IBM Corporation). The intrarater reliability was assessed using Cohen kappa. Paired *t* tests were performed to compare photo quality between the two models of cell phone (iPhone 4s and iPhone 6) and between before and after the photography education. Two-sample Student *t* tests were used to compare the photo quality between laypeople and dentists and between the photos taken of oneself and the subgroup partner.

Results

Each photo was assessed by the dentist twice. The kappa indexes for intrarater reliability were .764 and .728 for the general assessment and the assessments of hard tissue and soft tissue, respectively.

This study analyzed 965 photos: 441 taken by laypeople and 524 taken by dentists. Frontal view photos were taken by all subjects before and after receiving education. In contrast, occlusal area photos were taken by only 16 subjects (4 laypeople and 12 dentists) which were further taken by all subjects after education, and so these photos were not suitable for performing comparisons.

In the group of laypeople, photos in which the retraction was evaluated as being appropriate were taken by 8 subjects with the front camera and 12 with the rear camera by the partners before receiving education, increasing to 23 and 28, respectively, after receiving education. In the group of dentists, photos in which the retraction was evaluated as being appropriate were taken by 20 subjects with the front camera and 19 with the rear camera by the partners before receiving education, increasing to 28 and 30, respectively, after receiving education.

Photos taken after education had significantly higher scores than those taken before education with the exception of the evaluation item of optimal focus for the iPhone 6 front camera (Multimedia Appendix 1). Photos taken using the rear camera had significantly higher scores than those taken using the front camera by oneself (Multimedia Appendix 2). Although photos taken using the rear camera of the iPhone 6 had significantly higher scores in some items than those taken using the iPhone 4s, the iPhone 6 did not have overwhelming advantages (Multimedia Appendix 3). Photos taken by dentists had

significantly higher scores than those taken by laypeople for most of the evaluated items (Multimedia Appendix 4).

Discussion

Principal Results

This study assessed the use of cell phones for teleconsultations in dentistry. It was found that the photo scores were significantly higher for those taken by dentists than by laypeople for most of the evaluated items and for those taken after receiving education compared with beforehand except for one item in the iPhone 6 group with the front camera. The iPhone 6 did not have overwhelming advantages over the iPhone 4s.

There were common errors observed in the photos. Frontal view photos—and not occlusal view photos—were taken by both laypeople and dentists, and many images were taken without appropriate retraction. However, after participants received instruction in photography, the error rates in both the layperson and dentist groups were markedly reduced. Moreover, there were statistically significant differences in most of the evaluation categories between before and after receiving the education. These observations support the usefulness of providing image-taking instructions to both laypeople and medical staff in photography education sessions on appropriate methods for dental trauma teleconsultation using cell phones.

While the dentist group received higher evaluation scores both before and after the education compared with the laypeople group, there was no significant intergroup difference in the score for the item of shape in occlusal photos of hard tissue after the education using either the iPhone 4s or iPhone 6 ($P>.05$). These findings suggest that the laypeople can benefit from receiving education about how to take photos for use in dental evaluations, although not to the extent of dentists who have professional knowledge of dentistry. Comparisons based on the numbers of pixels of the camera showed that images taken with the rear camera scored higher than images taken with the front camera. For self-images taken using the front camera, images taken with the iPhone 4s (0.3 megapixels) had higher evaluation scores than images taken with the iPhone 6 (1.2 megapixels). This surprising result is probably due to factors other than the number of pixels, such as the size and weight of the device and the grip sensitivity. Moreover, the rear camera of a cell phone generally has a higher resolution than the front camera, and the rear cameras of both the iPhone 4s and iPhone 6 have 8.0 megapixels. However, the images taken with the iPhone 6 received higher scores probably because of its superior image sensor.

Comparison With Previous Work

According to the in vitro study of Park et al [26], autofocusing and white balance play important roles in photos taken using a cell phone for teleconsultation purposes. Similarly, our study found that focus, movement, and exposure characteristics affected the image quality. Another study using the same protocol as Park found that the precision of remote diagnoses was comparable to diagnoses conducted in person for photos that were taken by dentists [27].

There are several reports of satisfactory results being obtained when using cell phone apps for caries screening and traumatic dental injuries [25,28,29]. An intentionally developed cell phone app for traumatic dental injuries was recently evaluated for validity and usability [25]. That app allowed users to select from a library of images of the most common injuries and so could be used by any person regardless of their level of knowledge about dental trauma. However, its accuracy as well as the quality of the information reported under emergency conditions should be evaluated in the future.

In a study of the usefulness of an app in diagnosing dental trauma, Mohan et al [30] concluded that photos of injured teeth and soft tissue can be used by dentists to give a correct diagnosis. This indicates that a cell phone app that allows teleconsultation using the in-built cameras should be developed and used. The end goals of such a new app would be (1) to allow medical staff with minimal experience of dental trauma (ie, school nurses and emergency room medical staff) to send images of patients to dental trauma experts in real time and also receive a remote diagnosis and information about appropriate first aid treatment in real time and (2) for laypeople to be able to directly use the app and receive consultations to facilitate remote diagnosis and treatment via image transfer if a dentist is not available. In summary, the app should initially provide the basic functions of teleconsultation, and other additional functions should be theoretically grounded and evidence-based [31]. The International Association for Dental Traumatology provides a publicly available first aid protocol for dental trauma patients that patients can download from the internet to educate or treat themselves [32-34]. If an app that supports teleconsultation is developed, an additional protocol for performing an accurate diagnosis (ie, the imaging protocol suggested in our study) could be added to the instructions for patients.

However, a cell phone camera should be viewed as a complementary tool for use in teleconsultation, since images of dental trauma patients taken using cell phones may be different from images taken at a dental clinic. Moumoulidis et al [35] suggested that telemedicine does not always facilitate correct physician assessments, since their clinical trial found that 62% of diagnoses of nasal fractures based on images from cell phone cameras did not agree with the clinical assessments. This highlights that remote teleconsultations performed using cell phones do not guarantee accurate diagnoses, and so apps should be limited to use as a complementary tool to allow for appropriate emergency first aid in dental emergencies. Moreover, patients who receive teleconsultations should also seek appropriate treatment by visiting a dental clinic as soon as possible.

Limitations

This study was subject to several limitations. First, the study involved healthy subjects rather than actual trauma patients. Second, instead of evaluating images that had been taken with a cell phone and then transferred, the imaged files were analyzed directly on a computer screen. Third, we used iPhone 4s and iPhone 6 cell phones for the teleconsultations since most related studies have used iPhone devices, and so it might be worthwhile to repeat the experiment using Android cell phones in the future.

Future studies should address the limitations of this study including trying to mimic actual traumatic events.

Conclusions

Photography education is effective for both laypeople and dentists. Further developments of teleconsultation using the camera built into a cell phone require an optimal photo-taking protocol that should include the following factors: (1) obtaining frontal and occlusal images with proper retraction applied so that 6 anterior teeth and soft tissue are clearly visible, (2) photos should be taken by another person, and (3) the rear rather than the front camera should be used in order to optimize the image quality.

Acknowledgments

This research was supported by a grant from the Korea Health Technology Research and Development Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, Republic of Korea (grant number: HI18C0474).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparison of the quality of photos taken before and after education.
[DOCX File , 18 KB - mhealth_v8i3e15152_app1.docx]

Multimedia Appendix 2
Comparison of the quality of photos taken using front and rear cameras.
[DOCX File , 18 KB - mhealth_v8i3e15152_app2.docx]

Multimedia Appendix 3
Comparison of the quality of photos taken with the iPhone 4s and iPhone 6.
[DOCX File , 18 KB - mhealth_v8i3e15152_app3.docx]

Multimedia Appendix 4
Comparison of the quality of photos taken by laypeople and dentists.
[DOCX File , 18 KB - mhealth_v8i3e15152_app4.docx]

References

1. World Health Organization. 2018. Use of appropriate digital technologies for public health URL: http://apps.who.int/gb/ebwha/pdf_files/WHA71/A71_20-en.pdf [accessed 2019-11-11]
2. World Health Organization. 2016. Global diffusion of eHealth: making universal health coverage achievable URL: http://www.who.int/publications-detail/global-diffusion-of-ehealth-making-universal-health-coverage-achievable [accessed 2019-11-11]
3. Kim J, Kim S, Kim H, Kim K, Yang S, Lee C, et al. Effects of consumer-centered u-health service for the knowledge, skill, and attitude of the patients with chronic obstructive pulmonary disease. Comput Inform Nurs 2012 Dec;30(12):661-671. [doi: 10.1097/NXN.0b013e318261c1c] [Medline: 23266537]
4. Kim KM, Park KS, Lee HJ, Lee YH, Bae JS, Lee YJ, et al. Efficacy of a new medical information system, ubiquitous healthcare service with voice inception technique in elderly diabetic patients. Sci Rep 2015 Dec 11;5:18214 [FREE Full text] [doi: 10.1038/srep18214] [Medline: 26658492]
5. Wootton R. Recent advances: telemedicine. BMJ 2001 Sep 08;323(7312):557-560 [FREE Full text] [doi: 10.1136/bmj.323.7312.557] [Medline: 11546704]
6. Tachakra S, Wang XH, Istepanian RSH, Song YH. Mobile e-health: the unwired evolution of telemedicine. Telemed J E Health 2003;9(3):247-257. [doi: 10.1089/153056203322502632] [Medline: 14611692]
7. Johnston B, Wheeler L, Deuser J, Sousa KH. Outcomes of the Kaiser Permanente tele-home health research project. Arch Fam Med 2000 Jan;9(1):40-45. [Medline: 10664641]
8. Welch G, Balder A, Zagarins S. Telehealth program for type 2 diabetes: usability, satisfaction, and clinical usefulness in an urban community health center. Telemed J E Health 2015 May;21(5):395-403. [doi: 10.1089/tmj.2014.0069] [Medline: 25748544]
9. Blaivas M, Lyon M, Duggal S. Ultrasound image transmission via camera phones for overreading. Am J Emerg Med 2005 Jul;23(4):433-438. [Medline: 16032606]

10. Yamada M, Watarai H, Andou T, Sakai N. Emergency image transfer system through a mobile telephone in Japan: technical note. Neurosurgery 2003 Apr;52(4):986-988. [Medline: 12657199]
11. Goost H, Witten J, Heck A, Hadizadeh DR, Weber O, Gräff I, et al. Image and diagnosis quality of X-ray image transmission via cell phone camera: a project study evaluating quality and reliability. PLoS One 2012;7(10):e43402 [FREE Full text] [doi: 10.1371/journal.pone.0043402] [Medline: 23082108]
12. Chandhanayingyong C, Tangtrakulwanich B, Kiriratnikom T. Teleconsultation for emergency orthopaedic patients using the multimedia messaging service via mobile phones. J Telemed Telecare 2007;13(4):193-196. [doi: 10.1258/135763307780908049] [Medline: 17565775]
13. Stephens CD, Cook J. Attitudes of UK consultants to teledentistry as a means of providing orthodontic advice to dental practitioners and their patients. J Orthod 2002 Jun;29(2):137-142. [doi: 10.1093/ortho/29.2.137] [Medline: 12114464]
14. Yuen HK, Pope C. Oral home telecare for adults with tetraplegia: a feasibility study. Spec Care Dentist 2009;29(5):204-209. [doi: 10.1111/j.1754-4505.2009.00094.x] [Medline: 19740151]
15. Nickenig H, Wichmann M, Schlegel A, Eitner S. Use of telemedicine for pre-implant dental assessment: a comparative study. J Telemed Telecare 2008;14(2):93-97. [doi: 10.1258/jtt.2007.070806] [Medline: 18348756]
16. Lienert N, Zitzmann NU, Filippi A, Weiger R, Krastl G. Teledental consultations related to trauma in a Swiss telemedical center: a retrospective survey. Dent Traumatol 2010 Jun;26(3):223-227. [doi: 10.1111/j.1600-9657.2010.00873.x] [Medline: 20406276]
17. Gassner R, Tuli T, Hächl O, Moreira R, Ulmer H. Craniomaxillofacial trauma in children: a review of 3,385 cases with 6,060 injuries in 10 years. J Oral Maxillofac Surg 2004 Apr;62(4):399-407. [Medline: 15085503]
18. Andersson L, Malmgren B. The problem of dentoalveolar ankylosis and subsequent replacement resorption in the growing patient. Aust Endod J 1999 Aug;25(2):57-61. [Medline: 11411080]
19. Kopycka-Kedzierawski DT, Billings RJ. Teledentistry in inner-city child-care centres. J Telemed Telecare 2006;12(4):176-181. [doi: 10.1258/135763306777488744] [Medline: 16774697]
20. Estai M, Winters J, Kanagasingam Y, Shiikha J, Checker H, Kruger E, et al. Validity and reliability of remote dental screening by different oral health professionals using a store-and-forward telehealth model. Br Dent J 2016 Oct 07;221(7):411-414. [doi: 10.1038/sj.bdj.2016.733] [Medline: 27713449]
21. Aziz SR, Ziccardi VB. Telemedicine using smartphones for oral and maxillofacial surgery consultation, communication, and treatment planning. J Oral Maxillofac Surg 2009 Nov;67(11):2505-2509. [doi: 10.1016/j.joms.2009.03.015] [Medline: 19837324]
22. Hsieh C, Jeng S, Chen C, Yin J, Yang JC, Tsai H, et al. Teleconsultation with the mobile camera-phone in remote evaluation of replantation potential. J Trauma 2005 Jun;58(6):1208-1212. [Medline: 15995472]
23. Waran V, Selladurai BM, Bahuri NF, George GJ, Lim GP, Khine M. Teleconferencing using multimedia messaging service (MMS) for long-range consultation of patients with neurosurgical problems in an acute situation. J Trauma 2008 Feb;64(2):362-365. [doi: 10.1097/TA.0b013e318070cc88] [Medline: 18301199]
24. Park W, Kim D, Kim J, Kim K, Yoo SK. A portable dental image viewer using a mobile network to provide a tele-dental service. J Telemed Telecare 2009;15(3):145-149. [doi: 10.1258/jtt.2009.003013] [Medline: 19364900]
25. Zaror C, Espinoza-Espinoza G, Atala-Acevedo C, Muñoz-Millán P, Li Y, Clarke K, et al. Validation and usability of a mobile phone application for epidemiological surveillance of traumatic dental injuries. Dent Traumatol 2019 Feb;35(1):33-40. [doi: 10.1111/edt.12444] [Medline: 30248238]
26. Park W, Lee H, Jeong J, Kwon J, Lee GH, Kim K. Optimal protocol for teleconsultation with a cellular phone for dentoalveolar trauma: an in-vitro study. Imaging Sci Dent 2012 Jun;42(2):71-75 [FREE Full text] [doi: 10.5624/isd.2012.42.2.71] [Medline: 22783474]
27. de Almeida Geraldino R, Rezende LVML, da-Silva CQ, Almeida JCF. Remote diagnosis of traumatic dental injuries using digital photographs captured via a mobile phone. Dent Traumatol 2017 Oct;33(5):350-357. [doi: 10.1111/edt.12347] [Medline: 28453886]
28. Estai M, Kanagasingam Y, Xiao D, Vignarajan J, Bunt S, Kruger E, et al. End-user acceptance of a cloud-based teledentistry system and Android phone app for remote screening for oral diseases. J Telemed Telecare 2017 Jan;23(1):44-52. [doi: 10.1177/1357633X15621847] [Medline: 26721829]
29. AlShaya MS, Assery MK, Pani SC. Reliability of mobile phone teledentistry in dental diagnosis and treatment planning in mixed dentition. J Telemed Telecare 2018 Aug 22:1357633X18793767. [doi: 10.1177/1357633X18793767] [Medline: 30134778]
30. Mohan A, Agarwal T, Cherian TS, Muthu MS, Balasubramanian S, Subbalekshmi N, et al. Diagnostic ability of a smart phone app (injured tooth) in diagnosing traumatic injuries to the teeth: a multicentre analysis. Int J Paediatr Dent 2018 Nov;28(6):561-569. [doi: 10.1111/ipd.12411] [Medline: 29992696]
31. Tiffany B, Blasi P, Catz SL, McClure JB. Mobile apps for oral health promotion: content review and heuristic usability analysis. JMIR Mhealth Uhealth 2018 Sep 04;6(9):e11432 [FREE Full text] [doi: 10.2196/11432] [Medline: 30181114]
32. Diangelis AJ, Andreasen JO, Ebeleseder KA, Kenny DJ, Trope M, Sigurdsson A, International Association of Dental Traumatology. International Association of Dental Traumatology guidelines for the management of traumatic dental injuries:

1. Fractures and luxations of permanent teeth. Dent Traumatol 2012 Feb;28(1):2-12. [doi: 10.1111/j.1600-9657.2011.01103.x] [Medline: 22230724]
33. Andersson L, Andreasen JO, Day P, Heithersay G, Trope M, Diangelis AJ, International Association of Dental Traumatology. International Association of Dental Traumatology guidelines for the management of traumatic dental injuries: 2. Avulsion of permanent teeth. Dent Traumatol 2012 Apr;28(2):88-96. [doi: 10.1111/j.1600-9657.2012.01125.x] [Medline: 22409417]
34. Malmgren B, Andreasen JO, Flores MT, Robertson A, DiAngelis AJ, Andersson L, International Association of Dental Traumatology. International Association of Dental Traumatology guidelines for the management of traumatic dental injuries: 3. Injuries in the primary dentition. Dent Traumatol 2012 Jun;28(3):174-182. [doi: 10.1111/j.1600-9657.2012.01146.x] [Medline: 22583659]
35. Moumoulidis I, Mani N, Patel H, Leong P. A novel use of photo messaging in the assessment of nasal fractures. J Telemed Telecare 2007;13(8):387-390. [doi: 10.1258/135763307783064412] [Medline: 18078548]

Abbreviations

uHealth: ubiquitous health care
DSLR: digital single lens reflex

Edited by G Eysenbach; submitted 03.07.19; peer-reviewed by F Fatehi, A Garcia Linares; comments to author 20.10.19; revised version received 03.12.19; accepted 24.01.20; published 26.03.20.

Please cite as:
Jeong JS, Pang NS, Choi Y, Park KM, Kim T, Xu X, Park W
Importance of Photography Education to Improve Image Quality for Accurate Remote Diagnoses in Dental Trauma Patients: Observational Study
JMIR Mhealth Uhealth 2020;8(3):e15152
URL: http://mhealth.jmir.org/2020/3/e15152/
doi:10.2196/15152
PMID:32213475

Original Paper

Wound Care Knowledge, Attitudes, and Practices and Mobile Health Technology Use in the Home Environment: Cross-Sectional Survey of Social Network Users

Ya-Ting Kuan[1,2], RN, MSN; Tze-Fang Wang[2], RN, MSN; Chao-Yu Guo[3], PhD; Fu-In Tang[2], PhD; I-Ching Hou[2], PhD

[1]School of Nursing, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan

[2]School of Nursing, National Yang-Ming University, Taipei, Taiwan

[3]Institute of Public Health, National Yang-Ming University, Taipei, Taiwan

Corresponding Author:
I-Ching Hou, PhD
School of Nursing
National Yang-Ming University
Rm 407 Nursing Building
No 155, Sec 2, Li-Nong St, Beitou District
Taipei
Taiwan
Phone: 886 2 28267315
Email: evita@ym.edu.tw

Abstract

Background: Injury causing wounds is a frequent event. Inadequate or inappropriate treatment of injuries can threaten individual health. However, little is known about wound care knowledge, attitudes, and practices and mobile health (mHealth) use in the home environment in Taiwan.

Objective: This study aimed to evaluate wound care knowledge, attitudes, and practices and mHealth technology use among social network users.

Methods: A cross-sectional survey on social media platforms was conducted on adults aged 20 years and older. Data were collected from social network users in the home environment.

Results: A total of 361 participants were enrolled. The mHealth technology use of participants was positively correlated with wound care knowledge (r=.132, P=.01), attitudes (r=.239, P<.001), and practices (r=.132, P=.01). Participants did not have adequate knowledge (correct rate 69.1%) and were unfamiliar with the guidelines of proper wound care (correct rate 74.5%). Most participants had positive attitudes toward wound care and mHealth technology use. A total of 95.6% (345/361) of participants perceived that the use of mHealth technology can improve wound care outcomes, and 93.9% (339/361) perceived that wound care products should be optimized to be used with a mobile device. However, 93.6% (338/361) of participants had no experience using mHealth technology for wound care.

Conclusions: Our study shows the potential of mHealth technology to enhance wound care knowledge among social network users. Thus, government agencies and medical institutions in Taiwan should provide easy-to-use information products that enhance wound care knowledge, promote adequate behavior toward wound care, and prevent unpredictable or undesirable outcomes.

(JMIR Mhealth Uhealth 2020;8(3):e15678) doi:10.2196/15678

KEYWORDS

mobile health; wound; knowledge; attitudes; practices; home environment

Introduction

Injuries causing wounds occur frequently. In 2015, up to 50 million people worldwide incurred injuries because of road traffic crashes resulting in additional indirect health consequences associated with this growing epidemic [1]. Inadequate or inappropriate treatment of injuries can threaten individual health. Injuries contribute to approximately 10% of

mortality and 12% of morbidity worldwide [2]. In Taiwan, cases of road traffic accidents increased from 216,927 in 2007 to 403,906 in 2016, and this, in turn, resulted in an increase in the injury rate [3]. Similarly, in the United States, the number of road traffic accidents increased (from 2,491,000 in 2007 to 3,144,000 in 2016), and accordingly, the injury rate increased [4]. Wound care is one of the major challenges for health care systems [5,6] and accounts for 2% to 3% of the medical care budget [7]. Moreover, the demand for wound care is increasing. Patients who do not receive medical care or are discharged from medical institutions usually perform wound care by themselves at home. A study has found that 38.2% and 58.7% of patients returning from hospitals did not know how to change their dressing at home or how to clean the wounds, respectively [8]. Up to 84% of patients with surgical wounds return for regular follow-up [6]. If the wound is not properly treated, it may lead to infection (3% to 15%) [9,10]. Signs of wound infection include fever, swelling, pain, and purulent exudate. Factors such as corticosteroid use, smoking, and poor general health affect wound healing [11]. Patients who are not taught how to perform wound care and neglect the consequences of improper wound care experience a substantial economic burden and reduced quality of life [12].

Mobile health technology offers an alternative by increasing access to wound care resources [13,14] and involves the use of information and communication technology such as computers, mobile phones, personal digital assistants, and wearable sensors to deliver medical service and information [6,15]. In Taiwan, there were 29.31 million mobile communication users in 2019 [16]. The number of the population who are using a mobile phone to accessing internet is increase from 35.3% in 2011 to 88.2% in 2018 [17]. Mobile phone apps are an emerging tool for wound care [18]. Mobile health (mHealth) technology can support wound care knowledge, attitudes, and practices of social network users. Using mHealth technology (1) supports self-care among patients, (2) improves wound outcomes, (3) reduces care costs, (4) has built-in alerts, (5) enhances remote consultation, (6) promotes accurate assessment of wounds using wound images, (7) improves quality of life, and (8) supports nonspecialized caregivers [6,19-21]. Research has demonstrated that a lack of wound care knowledge and skills negatively affects the prognosis of wounds and health information technology use positively affects prognosis [8]. Users must be considered when designing health information technology [22]. Thus, wound care knowledge, attitudes, and practices and the mHealth technology use of social network users in Taiwan should be evaluated. However, little is known about these aspects. Therefore, this study aimed to evaluate the wound care knowledge, attitudes, and practices and mHealth technology use of individuals in the home environment in Taiwan. Findings regarding the use of mHealth technology among social network users were compared with prior work.

Methods

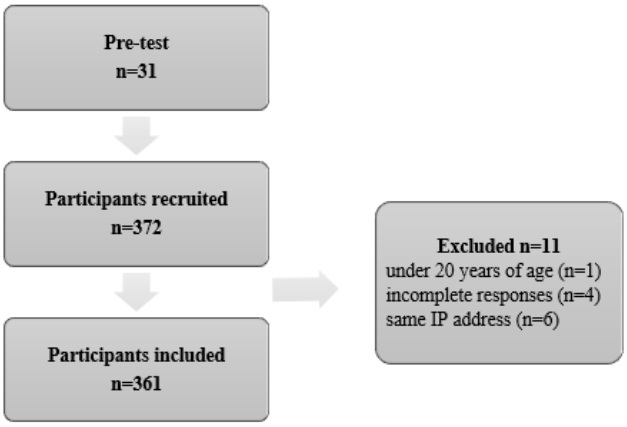
Study Design

We evaluated the wound care knowledge, attitudes, and practices and mHealth technology use of individuals in the home environment in Taiwan using a Web-based questionnaire survey. We conducted this Web-based questionnaire survey on a social media platform in the home environment between December 2015 and March 2016. Our subjects were Taiwanese individuals who had received wound care and sought health care information. A cross-sectional survey and network sampling were conducted to recruit subjects with background characteristics [23]. The study was approved by the National Yang-Ming University Human Research Ethics Committee (No YM104116E). A 4-part questionnaire was composed. Parts 1, 2, 3, and 4 were written to understand participant demographics, wound care experience, wound care competence, and mHealth technology use toward wounds, respectively. In the weeks preceding the formal investigation, a pretest was conducted in December 2015 with 31 subjects to ensure the clarity, conciseness, and readability of the scales and determine the approximate time required to complete the questionnaire. The formal investigation was anonymously conducted through a Web-based questionnaire survey addressed to participants in March 2016. Respondents were assured that their privacy was protected, and their informed consent was secured. To avoid duplicate responses, we only accepted the first response from the same internet protocol (IP) address without informing participants.

Participants and Recruitment

Participants aged 20 years and older who had experience using computers, communication devices, or consumer electronics (the “3C” products) and the ability to understand and complete an online informed consent form were included. Participants were recruited in March 2016 through social media platforms including Facebook and Professional Technology Temple (PTT) pages created specifically for the survey. PTT is one of the largest social media platforms in Taiwan [24]. Participants who did not meet the inclusion criteria and those who submitted multiple responses were excluded. According to an instrumental study, the sample size should be greater than 300 if the target population is over 5000 [25]. This survey was conducted from March 4 to March 31, 2016, and garnered 372 responses, with 11 responses excluded (1: aged younger than 20 years, 4: incomplete responses, and 6: same IP address). A total of 361 responses were included in the data analysis (Figure 1). The effective response rate was 97.0%.

Figure 1. Flow chart of the study design and screening process.



Instrument

We developed an instrument and defined questions according to the literature reviews [5,8,11]. With regard to questions related to wound care knowledge and wound care practices, correct-error scales (10=correct, 0=error) were used. Answers were considered correct if the items selected followed the wound care guidelines [5,8,11]; otherwise, answers were considered incorrect. With regard to questions related to wound care attitudes and mHealth technology use, arithmetic scales (10=strongly agree, 7.5=agree, 5=neutral, 2.5=disagree, 0=strongly disagree) were used.

Validity

To determine whether the instrument was appropriate and concise, we consulted wound care experts on content validity. Three experts—a licensed dermatologist working in a 2300-bed medical center; a nurse director at Taiwan Wound, Ostomy, and Continence Nurses Association; and an experienced nurse supervisor working in the burn center at a 1200-bed medical center—examined the entire instrument and offered suggestions and opinions on its content. These experts helped with appropriate wording and examined each item and the grouping. All the experts agreed that the four parts of the questionnaire were appropriate and clear. This questionnaire used a 4-point Likert scale with anchors ranging from 1=very inappropriate, 2=somewhat inappropriate, 3=appropriate, and 4=very appropriate and a 4-point Likert scale with anchors ranging from 1=very unclear, 2=somewhat unclear, 3=clear, and 4=very clear. The content validity index was 0.97, which was acceptable. After validation by the experts, the online instrument was created using Google Forms (Google LLC) and distributed to the enrolled participants.

Reliability

In the pilot study, 31 participants were invited to complete the online instrument; this was used to examine whether the wound care knowledge, attitudes, and practices items and mHealth technology use items in the instrument had internal consistency. Cronbach alpha for the study was 0.72, which indicated acceptable internal consistency reliability [26].

The survey consisted of approximately 60 questions and took between 15 and 20 minutes to complete. Participants were permitted to discontinue the survey anytime. Participants who completed the survey and provided an email address received an NT $100 (US $3.32) gift card to a retail store. Only completed questionnaires were included in the analysis. Participant identifiers were removed from the survey before data analysis.

Statistical Analysis

SPSS Statistics 23 software package (IBM Corporation) was used to analyze the data. Numbers, percentages, mean, standard deviation, and Pearson correlation (*r*) were used to examine the relationships between wound care knowledge, attitudes, and practices and mHealth technology use among the participants.

Results

Background Characteristics

A total of 361 social network users were included in this study. The main wound care knowledge resource among participants was health care professionals (367/1081, 34%), followed by experience/self-study (258/1081, 23.8%) and social media/internet/other (250/1081, 23.2%) (Table 1). Of participants, 60.7% (219/361) and 69.5% (251/361) had received medical treatment in the past 6 months and had incurred skin injuries in the previous year, respectively; 15.2% (55/361) were diagnosed with wound infection by a doctor; 24.4% (88/361) were not taught how to perform wound care; 33.2% (352/361) and 15.8% (342/361) used mobile phones and the internet, respectively; and while 78.1% (282/361) had been using mobile phones for more than 3 years, 93.6% (338/361) had never used the phone to look for wound care advice.

Table 1. Background characteristics of the participants (n=361).

Characteristic	Value
Gender, n (%)	
Male	235 (65.1)
Female	126 (34.9)
Age in years, n (%)	
20-24	156 (43.2)
25-29	114 (31.6)
30-34	59 (16.3)
≥35	32 (8.8)
Age in years, mean (SD)	26.5 (5.7)
Education level, n (%)	
Associate degree	15 (4.1)
Bachelor's degree	270 (74.9)
Master's/doctorate degree	76 (21.0)
Marital status, n (%)	
Married	37 (10.2)
Not married	324 (89.8)
Children, n (%)	
No	339 (93.9)
Yes	22 (6.1)
Wound care knowledge resource (select all that apply) (n=1081), n (%)	
School health education course	204 (18.9)
Health care professionals	367 (34.0)
Experience/self-study	258 (23.8)
Social media/internet/other	250 (23.3)

Wound Care Experience

The mean wound length and width were 4.02 (SD 5.71) cm and 2.40 (SD 2.73) cm, respectively. The primary causes of wounds were traffic accidents (124/361, 34.3%), penetrating injuries (59/361, 16.3%), and surgeries/diseases (58/361, 16.1%). A total of 48.5% (175/361) of participants incurred abrasion/contusion, 19.9% (72/361) incurred cuts, and 11.6% (42/361) other types of wounds. The primary causes of hospital visits were abrasions/contusions (184/606, 30.4%), lacerations (130/606, 21.5%), and cuts (105/606, 17.3%). Most of the participants incurred a wound on the knee (120/432, 27.8%), arm (87/432, 20.1%), or finger (86/432, 19.9%). Moreover, 89.8% (324/361) of participants applied the wound dressing by themselves, and 80.6% (291/361) had experience treating their own wound or treating others' wounds at the hospital. Many participants (260/361, 72.0%) were afraid or lacked confidence in treating their own or others' wounds. Wound care experience of the participants is presented in Table 2.

Table 2. Wound care experience of participants (n=361).

Characteristic	Value n (%)
Primary cause of wound	
Traffic accident	124 (34.3)
Penetrating injury	59 (16.3)
Surgery/disease	58 (16.1)
Falls	57 (15.8)
Other/unclear	38 (10.5)
Burn	13 (3.6)
Bite/scratch	12 (3.3)
Number of wounds	
1	226 (62.6)
2	72 (19.9)
3	31 (8.6)
>3	32 (9.1)
Type of wound	
Abrasion/contusion	175 (48.5)
Cuts	72 (19.9)
Other	42 (11.6)
Laceration	32 (8.9)
Unclear	24 (6.6)
Scratch	7 (1.9)
Insect bites	5 (1.4)
Bruising	4 (1.1)
Type of wound requiring hospital visit, select all that apply (n=606)	
Abrasion/contusion	184 (30.4)
Laceration	130 (21.5)
Cuts	105 (17.3)
Insect bites	54 (8.9)
Unclear	50 (8.3)
Bruising	47 (7.8)
Other	36 (6.0)
Location of wound, select all that apply (n=432)	
Knee	120 (27.8)
Arm	87 (20.1)
Finger	86 (19.9)
Leg	41 (9.5)
Other	31 (7.2)
Wrist	26 (6.0)
Buttock	18 (4.2)
Facial	13 (3.0)
Head	10 (2.3)
Wound appearance	
No sign of infection	122 (19.4)

Characteristic	Value n (%)
Dirty	97 (15.4)
Partial thickness skin loss	181 (28.8)
<10 cc bleeding	201 (32.0)
>10 cc bleeding	28 (4.5)
Type of wound disinfectant used, select all that apply (n=550)	
Povidone-iodine solution	242 (44.0)
Normal saline solution	200 (36.4)
Antibiotic ointment	74 (13.5)
Mercurochrome/acrinol	23 (4.1)
Other	11 (2.0)
Type of wound dressing used, select all that apply (n=534)	
Gauze	216 (40.4)
Adhesive bandage	135 (25.3)
DuoDERM	107 (20.0)
Antimicrobial dressing	30 (5.6)
No dressing used	27 (5.1)
Collagen dressing	6 (1.1)
Chinese medicine dressing	6 (1.1)
Other	7 (1.3)

Wound Care Knowledge, Attitudes, and Practices

For wound care knowledge, more participants could correctly identify a photo of an abrasion wound than a diabetic foot or pressure ulcer wound. The percentage of participants with wound care knowledge was 69.1% (Table 3).

For wound care practices, most participants assessed wound appearance before dressing (339/361, 93.9%). Most participants washed their hands before the last wound dressing they performed (327/361, 90.6%). Less than one-quarter of participants (84/361, 22.4%) used a sterile cotton swab for wound dressing. The mean rate of correct wound care practices was 74.5% (Table 3).

For the wound care attitudes, 28.5% (103/361) of participants showed that they had good knowledge of assessing a wound (Table 4). Half of participants (183/361, 50.7%) worried about lacking ability to perceive wound infection or complication. Most participants (330/361, 91.5%) thought that the method of managing wounds was important for wound healing. However, only 27.1% (98/361) of participants had the confidence to care for wounds correctly.

Table 3. Wound care knowledge and practices (n=361).

Topic	Score[a] (Correct %)	Rank
Wound care knowledge		
Identify image of an abrasion wound	9.9 (86.4[b])	1
Identify image of a diabetic foot	6.4 (63.7[c])	2
Identify image of a pressure ulcer wound	4.3 (42.7[d])	3
I believe not smoking promotes wound healing	10 (99.7)	1
I believe nutrition may be a factor in promoting wound healing	9.4 (94.2)	2
I believe getting enough sleep may be a factor in promoting wound healing	9.3 (93.4)	3
I believe not using steroids may be a factor in promoting wound healing	7.6 (75.6)	4
I believe keeping the moisture balance of the wound bed can help wound healing	5.8 (58.0)	5
I believe appropriate exercise may be a factor in promoting wound healing	4.0 (40.1)	6
I believe abnormal exudate may be a sign of wound infection	9.4 (93.6)	1
I believe redness and swelling may be signs of wound infection	7.8 (77.6)	2
I believe fever may be a sign of wound infection	6.8 (67.6)	3
I believe pain may be a sign of wound infection	6.3 (63.4)	4
I believe cold may be a sign of wound infection	1.2 (11.6)	5
Wound care practices		
Assess wound appearance before dressing (eg, redness, exudate)	9.4 (93.9)	1
Wash hands before dressing wound	9.1 (90.6)	3
Remove gauze after rinsing wound with normal saline solution or boiled water	6.7 (67.1[e])	7
Use normal saline solution or boiled water to clean wound	7.0 (69.6[f])	6
Use sterile cotton swab to dress wound	2.2 (22.4[g])	8
Use dressing that covers wound margin by at least 1 cm all around	9.3 (92.8[h])	2
Contact position between the finger/clip and the dressing when covering	7.7 (76.8[i])	5
Wash hands after dressing wound	8.3 (82.9)	4

[a]Correct-error scales: 10=correct, 0=error.

[b]Bruising (0.8%), laceration (10.2%), cuts (0.3%), burns (1.9%), arteriovenous ulcer (0.3%).

[c]Bruising (2.5%), laceration (0.3%), abrasion/contusion (1.1%), burns (20.5%), arteriovenous ulcer (6.1%), pressure ulcer (5.8%).

[d]Bruising (3.6%), laceration (5.5%), cuts (0.6%), abrasion/contusion (10%), burns (13.6%), diabetic foot (17.7%), arteriovenous ulcer (6.4%).

[e]Removing sticking gauze directly (23.0%), removing gauze after rinsing with tap water (2.5%), removing gauze after rinsing with povidone-iodine solution (2.5%), removing gauze after rinsing with alcohol-iodine solution (2.8%), removing gauze after rinsing with hydrogen peroxide solution (1.1%), other (1%).

[f]No cleaning of wound (5.8%), tissue (7.2%), gauze (5.3%), tap water (10.2%), other (1.7%).

[g]Tissue (3.6%), gauze (15.5%), nonsterile cotton swab (22.4%), other (1.1%).

[h]Smaller than wound's margin (2.2%); equal to wound's margin (5%).

[i]The wound contact side of dressing (23.2%).

Table 4. Wound care attitudes (n=361).

Attitude	Negative (score 0, 2.5[a]) n (%)	Neutral (score 5[a]) n (%)	Positive (score 7.5, 10[a]) n (%)
I know very well how to assess wound.	63 (17.5)	195 (54.0)	103 (28.5)
I am worried about my lack of ability to perceive wound infection or complication.	183 (50.7)	110 (30.5)	68 (18.8)
I think that the method of managing a wound is important for wound healing.	6 (1.7)	25 (6.9)	330 (91.4)
I am confident I am doing wound care correctly.	77 (21.3)	186 (51.5)	98 (27.2)

[a]Arithmetic scale: 10=strongly agree, 7.5=agree, 5=neutral, 2.5=disagree, 0=strongly disagree.

Mobile Health Technology Use

Most of the participants responded neutral to positively to the following (Table 5): it is important to use mHealth technology in wound care (345/361, 95.6%); the use of mHealth technology in wound care can be helpful in improving wound care outcomes (345/361, 95.6%); wound care information products should be optimized for mobile devices (339/361, 93.9%) and should be easy to use (341/361, 94.4%); and I am interested in how mHealth technology can help me take care of wounds (347/361, 96.1%). The mHealth technology use of participants was positively correlated with wound care knowledge (r=.132, P=.01), attitudes (r=.239, P<.001), and practices (r=.132, P=.01).

Table 5. Mobile health technology use of participants (n=361).

Content	Negative (score 0, 2.5[a]) n (%)	Neutral (score 5[a]) n (%)	Positive (score 7.5, 10[a]) n (%)
It is important to use mobile health technology in wound care.	16 (4.4)	164 (45.4)	181 (50.2)
The use of mobile health technology in wound care can be helpful in improving wound care outcomes.	16 (4.4)	109 (30.2)	236 (65.4)
Wound care information products should be optimized for mobile devices.	22 (6.1)	129 (35.7)	210 (58.2)
Wound care information products should be easy to use.	20 (5.6)	99 (27.4)	242 (67.0)
I am interested in how mobile health technology can help me take care of wounds.	14 (3.9)	92 (25.5)	255 (70.6)

[a]Arithmetic scale: 10=strongly agree, 7.5=agree, 5=neutral, 2.5=disagree, 0=strongly disagree.

Discussion

Principal Findings

This study showed that most participants (345/361, 95.6%) understood the importance of using mHealth technology and 96.1% (347/361) of participants were interested in using this method to help them take care of wounds at home. Most participants (324/361, 89.8%) had wound dressing experience, but participants lacked confidence in their ability to assess and perform wound care correctly (260/361, 72%). Most wounds can be self-managed at home [27]. Hence, patients should be taught how to perform basic wound assessment and determine any signs of complications and infections [8]. Most participants had a positive attitude toward wound care and mHealth technology use, but they did not have adequate knowledge (correct rate 69.1%) and were unfamiliar with the guidelines in performing proper wound care (correct rate 74.5%). Removing gauze sticking to the wound without wetting the gauze, which might harm granulating tissue, was not commonly done. A total of 30.4% of the participants did not use normal saline solution or boiled, sterilized water in cleaning the wound. The use of disinfectants to clean wounds with no sign of infection is not conducive to wound healing; hence, cleaning uninfected wounds with normal saline or boiled, sterilized water was suggested [11]. In addition, the sterile concept should be explained since patients used nonsterile cotton swabs, easily available at home, to clean wounds. Additionally, whether nonsterile cotton swabs can be used as an alternative to sterile cotton swabs in nonhospital settings should be further studied.

Comparison With Prior Work

Wound care education can improve wound care attitudes and reduce the fear of taking care of wounds [5,10]. Participant wound care knowledge was positively correlated with wound care practices [28,29]. Thus, improving the wound care knowledge of the patient can enhance wound care skills [5]. Participants' main resources for wound care knowledge were health care professionals. Other studies have found similar trends in which most social network users responded that contact with online professionals was somewhat important or very important [30]. Our study had similar results; in addition, we showed that participant wound care competency was positively correlated with mHealth technology use. Thus, policymakers should focus on increasing mHealth technology use and wound care competency among social network users in Taiwan. In addition, 95.6% of participants presumed that the application of mHealth technology can improve wound care outcomes, and 93.9% (339/361) perceived that wound care information products should be optimized to use with a mobile device. Developing software for mobile devices such as visual reality simulation for wound care to increase the interaction with wound care learners should be considered. In addition, developing artificial intelligence that can determine wound type or signs

of infection is recommended, and a health care professional must be consulted to receive the correct care recommendations.

Limitations

This study had some limitations. A single platform was used in selecting the sample. The findings of this study are limited to adults living in Taiwan, and the mean age was 26.5 years. Samples were not representative of all social network users in Taiwan as there were restrictions in patients' age and wound type. Moreover, as self-report was adopted to understand the wound care knowledge, attitudes, and practices, it is possible that participants adhered to perceived social norms.

Conclusion

Our study showed the potential of mHealth technology in enhancing wound care knowledge, attitudes, and practices among social network users in Taiwan. Most participants responded that it is important to apply mHealth technology in wound care. However, most of them had not used mHealth technology for wound care. Therefore, our results can serve as an important reference for conducting further studies on the use of mHealth technology in wound care among social network users in the home environment. The association among wound care knowledge, attitudes, and practices and mHealth technology suggests that government agencies and medical institutions should provide correct information for wound care knowledge to promote appropriate behavior toward wound care and prevent unpredictable or undesirable outcomes.

Acknowledgments

We are grateful to the participants of this study. Supported by grant from Yen Tjing Ling Medical Foundation (project number CI-107-34).

Authors' Contributions

All authors made substantial contributions to the conception and design of the study, acquisition of data, and analysis or interpretation of data. All authors revised and approved the manuscript.

Conflicts of Interest

None declared.

References

1. World Health Organization. 2015. Global status report on road safety 2015 URL: http://apps.who.int/iris/bitstream/10665/189242/1/9789241565066_eng.pdf?ua=1 [accessed 2020-01-22]
2. Bhalla K, Harrison JE, Shahraz S, Fingerhut LA, Global Burden of Disease Injury Expert Group. Availability and quality of cause-of-death data for estimating the global burden of injuries. Bull World Health Organ 2010 Nov 01;88(11):831-838C [FREE Full text] [doi: 10.2471/BLT.09.068809] [Medline: 21076564]
3. Ministry of Transportation and Communications (MOTC). [Selected indicators of transportation and communications] URL: http://www.motc.gov.tw/ch/home.jsp?id=59&parentpath=0,6 [accessed 2020-01-22]
4. National Highway Traffic Safety Administration. 2016. Summary of motor vehicle crashes URL: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812580 [accessed 2020-01-22]
5. Chen Y, Wang Y, Chen W, Smith M, Huang H, Huang L. The effectiveness of a health education intervention on self-care of traumatic wounds. J Clin Nurs 2012 Nov 02;22(17-18):2499-2508. [doi: 10.1111/j.1365-2702.2012.04295.x]
6. Morris MS, Deierhoi RJ, Richman JS, Altom LK, Hawn MT. The relationship between timing of surgical complications and hospital readmission. JAMA Surg 2014 Apr;149(4):348-354. [doi: 10.1001/jamasurg.2013.4064] [Medline: 24522747]
7. Edwards H, Finlayson K, Courtney M, Graves N, Gibb M, Parker C. Health service pathways for patients with chronic leg ulcers: identifying effective pathways for facilitation of evidence based wound care. BMC Health Serv Res 2013 Mar 8;13(1):86. [doi: 10.1186/1472-6963-13-86]
8. Pieper B, Sieggreen M, Nordstrom CK, Freeland B, Kulwicki P, Frattaroli M, et al. Discharge knowledge and concerns of patients going home with a wound. J Wound Ostomy Continence Nurs 2007;34(3):245-253. [doi: 10.1097/01.WON.0000270817.06942.00] [Medline: 17505242]
9. Quinn JV, Polevoi SK, Kohn MA. Traumatic lacerations: what are the risks for infection and has the 'golden period' of laceration care disappeared? Emerg Med J 2014 Feb;31(2):96-100 [FREE Full text] [doi: 10.1136/emermed-2012-202143] [Medline: 23314208]
10. Pieper B. Vulnerable populations: considerations for wound care. Ostomy Wound Manage 2009 May 01;55(5):24-37 [FREE Full text] [Medline: 19471046]
11. Lewis S, Dirksen S, Heitkemper M, Bucher L, Harding M. Medical Surgical Nursing: Assessment and Management of Clinical Problem. St. Louis: Mosby; 2014.

12. Tricco AC, Antony J, Vafaei A, Khan PA, Harrington A, Cogo E, et al. Seeking effective interventions to treat complex wounds: an overview of systematic reviews. BMC Med 2015 Apr 22;13:89 [FREE Full text] [doi: 10.1186/s12916-015-0288-5] [Medline: 25899006]
13. Scott AR, Alore EA, Naik AD, Berger DH, Suliburk JW. Mixed-methods analysis of factors impacting use of a postoperative mhealth app. JMIR Mhealth Uhealth 2017 Feb 08;5(2):e11 [FREE Full text] [doi: 10.2196/mhealth.6728] [Medline: 28179215]
14. Lee E, Han S, Jo SH. Consumer choice of on-demand mHealth app services: context and contents values using structural equation modeling. Int J Med Inform 2017 Dec;97:229-238. [doi: 10.1016/j.ijmedinf.2016.10.016] [Medline: 27919381]
15. Brian RM, Ben-Zeev D. Mobile health (mHealth) for mental health in Asia: objectives, strategies, and limitations. Asian J Psychiatr 2014 Aug;10:96-100. [doi: 10.1016/j.ajp.2014.04.006] [Medline: 25042960]
16. Taiwan National Communications Commission. [Statistics for the mobile communications market for the third quarter of 2019] URL: https://www.ncc.gov.tw/chinese/files/19122/3773_42434_191226_1.pdf [accessed 2020-01-29]
17. Taiwan National Development Council. [Survey of the population who have mobile phones in 2018] URL: https://tinyurl.com/vpfsok7 [accessed 2020-01-29]
18. Chao E, Meenan CK, Ferris LK. Smartphone-based applications for skin monitoring and melanoma detection. Dermatol Clin 2017 Oct;35(4):551-557. [doi: 10.1016/j.det.2017.06.014] [Medline: 28886812]
19. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE. Development of mHealth system for supporting self-management and remote consultation of skincare. BMC Med Inform Decis Mak 2015 Dec 30;15:114 [FREE Full text] [doi: 10.1186/s12911-015-0237-4] [Medline: 26714452]
20. Silva BMC, Rodrigues JJPC, de la Torre Diez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272. [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
21. Onodera R, Sengoku S. Innovation process of mHealth: an overview of FDA-approved mobile medical applications. Int J Med Inform 2018 Oct;118:65-71. [doi: 10.1016/j.ijmedinf.2018.07.004]
22. Tunnell H, Faiola A, Bolchini D, Bartlett Ellis R. Simulated clinical encounters using patient-operated mhealth: experimental study to investigate patient-provider communication. JMIR Mhealth Uhealth 2018 Nov 01;6(11):e11131 [FREE Full text] [doi: 10.2196/11131] [Medline: 30389652]
23. Grove S, Burns N, Gray J. The Practices of Nursing Research: Appraisal, Synthesis. St. Louis: Elsevier Health Sciences; 2008.
24. Wikipedia. PTT URL: https://zh.wikipedia.org/wiki/%E6%89%B9%E8%B8%A2%E8%B8%A2 [accessed 2020-01-22]
25. Hair J, Black W, Babin B, Anderson R. Multivariate Data Analysis: A Global Perspective. Upper Sadder River: Prentice-Hall; 2010.
26. Gay L. Educational Research: Competencies for Analysis and Applications. New York: Longman; 1992.
27. Lin M, Lee H, Yu W, Wang S, Lin M. The epidemiology of illness and injury during the 2010 Taipei International Flora Exposition. Taiwan J Public Health 2013;32(5):502-515. [doi: 10.6288/TJPH2013-32-05-09]
28. Eng JJ, Noonan VK, Townson AF, Higgins CE, Rogers J, Wolfe DL. Impact of an online medical internet site on knowledge and practice of health care providers: a mixed methods study of the Spinal Cord Injury Rehabilitation Evidence project. J Med Internet Res 2014 Dec 23;16(12):e296 [FREE Full text] [doi: 10.2196/jmir.3453] [Medline: 25537167]
29. Jordan DN, Jordan JL. Foot self-care practices among Filipino American women with type 2 diabetes mellitus. Diabetes Ther 2011 Mar;2(1):1-8 [FREE Full text] [doi: 10.1007/s13300-010-0016-2] [Medline: 22127764]
30. Wetterlin FM, Mar MY, Neilson EK, Werker GR, Krausz M. eMental health experiences and expectations: a survey of youths' Web-based resource preferences in Canada. J Med Internet Res 2014;16(12):e293 [FREE Full text] [doi: 10.2196/jmir.3526] [Medline: 25519847]

Abbreviations

IP: internet protocol
mHealth: mobile health
PTT: Professional Technology Temple

Edited by G Eysenbach; submitted 29.07.19; peer-reviewed by C Jacob, K Saleem; comments to author 03.10.19; revised version received 29.10.19; accepted 16.12.19; published 26.03.20.

Please cite as:
Kuan YT, Wang TF, Guo CY, Tang FI, Hou IC
Wound Care Knowledge, Attitudes, and Practices and Mobile Health Technology Use in the Home Environment: Cross-Sectional Survey of Social Network Users
JMIR Mhealth Uhealth 2020;8(3):e15678
URL: http://mhealth.jmir.org/2020/3/e15678/
doi:10.2196/15678
PMID:32213478

Original Paper

Leveraging Providers' Preferences to Customize Instructional Content in Information and Communications Technology–Based Training Interventions: Retrospective Analysis of a Mobile Phone–Based Intervention in India

Hanu Tyagi[1,2], BTECH, MBA; Manisha Sabharwal[3], MSc; Nishi Dixit[3], BTECH; Arnab Pal[3], BTECH, MBA; Sarang Deo[2], MBA, PhD

[1]Carlson School of Management, University of Minnesota, Minneapolis, MN, United States
[2]Max Institute of Healthcare Management, Indian School of Business, Hyderabad, India
[3]Clinton Health Access Initiative, New Delhi, India

Corresponding Author:
Hanu Tyagi, BTECH, MBA
Carlson School of Management
University of Minnesota
321 19th Ave S
Minneapolis, MN, 55455
United States
Phone: 1 7637427688
Email: tyagi035@umn.edu

Abstract

Background: Many public health programs and interventions across the world increasingly rely on using information and communications technology (ICT) tools to train and sensitize health professionals. However, the effects of such programs on provider knowledge, practice, and patient health outcomes have been inconsistent. One of the reasons for the varied effectiveness of these programs is the low and varying levels of provider engagement, which, in turn, could be because of the form and mode of content used. Tailoring instructional content could improve engagement, but it is expensive and logistically demanding to do so with traditional training

Objective: This study aimed to discover preferences among providers on the form (articles or videos), mode (featuring peers or experts), and length (short or long) of the instructional content; to quantify the extent to which differences in these preferences can explain variation in provider engagement with ICT-based training interventions; and to compare the power of content preferences to explain provider engagement against that of demographic variables.

Methods: We used data from a mobile phone–based intervention focused on improving tuberculosis diagnostic practices among 24,949 private providers from 5 specialties and 1734 cities over 1 year. Engagement time was used as the primary outcome to assess provider engagement. K-means clustering was used to segment providers based on the proportion of engagement time spent on content formats, modes, and lengths to discover their content preferences. The identified clusters were used to predict engagement time using a linear regression model. Subsequently, we compared the accuracy of the cluster-based prediction model with one based on demographic variables of providers (eg, specialty and geographic location).

Results: The average engagement time across all providers was 7.5 min (median 0, IQR 0-1.58). A total of 69.75% (17,401/24,949) of providers did not consume any content. The average engagement time for providers with nonzero engagement time was 24.8 min (median 4.9, IQR 2.2-10.1). We identified 4 clusters of providers with distinct preferences for form, mode, and length of content. These clusters explained a substantially higher proportion of the variation in engagement time compared with demographic variables (32.9% vs 1.0%) and yielded a more accurate prediction for the engagement time (root mean square error: 4.29 vs 5.21 and mean absolute error: 3.30 vs 4.26).

Conclusions: Providers participating in a mobile phone–based digital campaign have inherent preferences for instructional content. Targeting providers based on individual content preferences could result in higher provider engagement as compared to targeting providers based on demographic variables.

(JMIR Mhealth Uhealth 2020;8(3):e15998) doi:10.2196/15998

KEYWORDS

public health; mobile health; health care providers; health care workers; instructional technology; information technology; infectious diseases; provider training; learning preferences

Introduction

The recent proliferation and adoption of information and communications technology (ICT) have the potential to transform learning among health professionals [1]. Riding the technology wave, many public health programs and interventions across geographies and therapeutic areas leverage ICT-based interventions to train and sensitize health professionals [2]. They provide a cost-effective mechanism to reach professionals [3], especially to engage with geographically distant or fragmented providers [4].

However, the evidence regarding the effects of ICT-based interventions on provider knowledge, attitude, practice, and, consequently, patient health outcomes is not unequivocal [5]. One of the reasons for their uncertain effectiveness is low [6] and varying levels [7] of provider engagement within the ICT-based interventions. Among various factors that could explain heterogeneity [8], the format and mode of instructional content are known to play an important role in improving provider engagement with ICT-based interventions [9,10].

Prior research has shown that the customization of instructional content could enhance the learning experience in academic settings among medical students [11]. In the case of nonacademic provider-focused training interventions as well, customization of instructional content could improve provider engagement, but it may not always be feasible. Especially, in the case of traditional training methods such as lectures and conferences, the assessment of learners' preferences could be expensive, and customized delivery of content could be logistically demanding [12]. In contrast, ICT-based training interventions make it feasible to tailor content to providers' preferences, providing the much-needed learner-centric approach [13].

This study aimed to determine content preferences among providers in terms of form (articles or videos), mode (featuring peers or experts), and length (short or long) of instructional content. We used the inherent content preferences among providers to explain variation in provider engagement with ICT-based training interventions. We compared the magnitude of variation in provider engagement explained by content preferences with that explained by demographic variables. The research questions have been addressed by analyzing data from a mobile phone–based provider training intervention, which was designed to improve tuberculosis (TB) diagnostic practices among private providers in India.

Methods

Study Setting and Participants

We collaborated with a third-party mobile phone–based platform, which helps providers to discuss real-life medical cases with the provider community. At the start of the intervention, more than 225,000 private providers were registered on the platform. For our study, we chose providers who logged into the platform at least once a month and further narrowed our list by choosing providers from 5 specialties—general practice, internal medicine, obstetrics and gynecology, pediatrics, and pulmonology—which account for the bulk of patients with TB initiated on anti-TB treatment [14]. A total of 24,949 private providers spread across 1734 cities and towns in India participated in the campaign, which ran from February to November 2017.

Intervention

We launched a digital campaign on the mobile phone–based platform through a dedicated page called *ThinkTB*. The digital campaign focused on the dissemination of TB diagnostic best practices among private providers in India. It showcased 10 content pieces—5 videos, 3 articles, and 2 interactive games (Table 1). In the first phase, the campaign content aimed at raising awareness and interest among providers. In the second phase, it progressed to interactive content and games to inculcate trial and advocacy. We tailored the content pieces according to the providers' specialty (Figure 1) to highlight their respective roles in the diagnosis of TB. For example, for video 2 launched in April 2017, the pediatricians were shown a video titled *Role of pediatricians in diagnosing TB*, whereas the gynecologists were shown a video titled *Female genital tract TB: A diagnosing challenge?*

Table 1. Details of content pieces delivered in the ThinkTB campaign.

Month and content	Topic	Target specialty
February 2017		
Video 1	• Role of GP[a] in the management of TB[b] in India • Keeping up with the changing diagnostic paradigms in TB • Inability to conceive: could this be TB? • Role of GP in the management of TB in India	• General practice • Internal medicine and pulmonology • Gynecology • General practice
March 2017		
Article 1	• Role of GPs in the management of TB in India • Vague presentation of TB in pediatric population • Inability to conceive: could this be TB? • Comparative features of tests for diagnosis of tuberculosis	• General practice • Pediatrics • Obstetrics and gynecology • Internal medicine and pulmonology
Webcast 1	• Role of GPs in diagnosis of TB • Female genital TB: myths and facts • Role of pulmonologists in diagnosing TB • Endorsed tests for diagnosis of pulmonary and extra-pulmonary TB	• General practice • Gynecology • Pulmonologists • Internal medicine
April 2017		
Video 2	• Role of pediatricians in diagnosing TB • Female genital tract TB: a diagnosing challenge? • Tuberculosis: a growing health concern • Tuberculosis: guide to early detection	• Pediatrics • Gynecology • General practice • Internal medicine and pulmonology
Article 2	• Female genital tuberculosis: a diagnosing challenge? • Tuberculosis: a growing health concern • Pediatric tuberculosis: an overview • Tuberculosis: all you need to know • Tuberculosis: guide to early detection	• Gynecology • General practice • Pediatrics • Pulmonology • Internal medicine
June 2017		
Article Gyn	• An article on drug resistant tuberculosis (10 principles for effective management)	Gynecology
Calculator	• Efficient diagnostic tool	All
September 2017		
Expert video 1[c]	• How does one diagnose and treat MDR-TB[d]? • How does one treat tuberculosis? • How to interpret discordant results? • Complex case of tuberculosis • Complex case of FGTB[e] • Complex case of drug resistance	All
October 2017		
BMJ training	• Accredited E-training module extrapulmonary and pulmonary tuberculosis	All
November 2017		
Expert video 2[c]	• What are the recommended tests for pulmonary tuberculosis and which tests are discouraged? • What test can be used for diagnosing tuberculosis pleural effusion? • What is the ideal or the best diagnostic algorithm for tuberculosis today?	All

[a]GP: general practitioner.

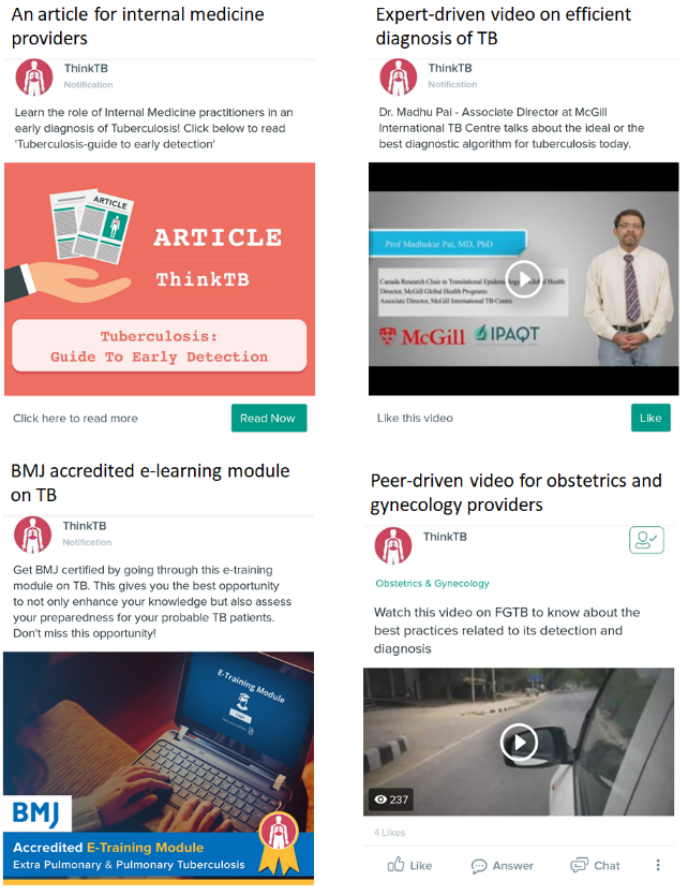
[b]TB: tuberculosis.

[c]The mapping of expert videos with the specialty was 1 to many.

[d]MDR-TB: multidrug-resistant tuberculosis.

[e]FGTB: female genital tuberculosis.

Figure 1. Mobile screenshots of content pieces from the ThinkTB campaign.



Data

We had access to 3 datasets pertaining to the providers, content pieces, and the interaction of the provider with the content pieces. The first dataset was unique at a provider level and included provider information—provider specialty and city of practice—which was recorded when the providers registered on the mobile platform. The third-party platform had segmented cities into 3 *city tiers* (tier 1, tier 2, and tier 3) based on a classification used by the Indian government to set the minimum daily wage for agricultural and industrial workers in India and estimate the house rent allowance provided by the Central Government of India to its employees [15]. The second dataset was unique at the content piece level and included classification of each content piece by its format (articles or videos), mode (videos featuring either peers or subject matter experts from the provider community), and length (long content, which is >10 min, or short content, which is ≤10 min) as shown in Table 2. The third dataset was unique at the provider-content piece level and recorded the engagement time defined as the time spent by a provider on a content piece (reading or viewing). We defined the content piece to be *consumed* if the engagement time associated with the content piece was greater than zero.

We combined the 3 datasets to obtain a comprehensive dataset with 24,949 observations, each representing a unique provider. It contained the following variables: provider specialty; city; city tier; the engagement time spent on each content piece; and the format, mode, and length of the content pieces. We excluded providers who had zero engagement time and further removed outliers with engagement time greater than the sum of the third quartile and 1.5 times the IQR. The remaining providers with nonzero engagement time were taken into consideration for analysis.

Table 2. Categorization of content pieces.

Content piece	Format	Length	Mode[a]
Article 1	Article	Short	N/A[b]
Video 1	Video	Short	Peer
Webcast 1	Video	Long	Peer
Article 2	Article	Short	N/A
Video 2	Video	Short	Peer
Calculator	Article	Short	N/A
Article Gyn	Article	Short	N/A
Expert video 1	Video	Short	Expert
Expert video 2	Video	Short	Expert
BMJ training	Article	Long	N/A

[a]Mode is only defined for videos.

[b]Not applicable.

Analysis

We used engagement time with the campaign as the primary outcome to assess provider engagement. We conducted our analysis in 2 steps as has been described in the following subsections. As described in the subsection *Discovering Content Preferences Among Providers*, providers were segmented into clusters based on the proportion of time spent on content formats and modes to discover their content preferences. As explained in the subsection *Predicting Provider Engagement*, separate regression models were fitted to assess the extent to which the variation in engagement time can be explained by clusters and demographic variables. All analyses were performed using R 3.4.3 (The R Foundation) [16].

Discovering Content Preferences Among Providers

We calculated 3 proportions of engagement time spent by a provider using the content classifications described earlier. For the 2 formats (text and video), *read proportion* was calculated as the proportion of engagement time spent by a provider on reading articles. For the 2 modes of videos (peer and expert), *expert proportion* was calculated as the proportion of engagement time spent by a provider on videos featuring an expert. For the 2 types of lengths, *short proportion* was calculated as the proportion of engagement time spent on consuming short content pieces.

These 3 proportions were used to cluster providers using the k-means algorithm [17]. The k-means algorithm partitioned providers into clusters such that the providers were well matched to other providers in their own clusters but were very different from those in the other clusters. The optimal number of clusters was determined using the average silhouette width method [18]. A higher value of silhouette width of observation indicates that the observation is well matched to its cluster and poorly matched to neighboring clusters. We varied the number of clusters from 2 to 15 to determine an optimal number of clusters. Moreover, Chi-square tests were performed to assess the similarity of clusters in terms of the city tier and specialty distribution among them.

Predicting Provider Engagement

We estimated 2 linear regression models to explain the variation in provider engagement with the campaign using engagement time as the outcome variable. For the first model (model A), clusters were used as a predictor variable. For the second model (model B), provider-level demographic variables (specialty and city tier) were used as predictor variables. The number of providers in the regression models was 6482 after data cleaning, as has been explained in previous sections.

The demographic variable–based model was compared with the cluster-based model based on its predictive power. A 10-fold cross-validation method was used, which is commonly used for model selection [19]. For each dataset, 9 parts were used for training the model and the tenth part for testing it. This process was repeated 10 times, ensuring that each dataset partition served as a test set. We then compared the average of cross-validation root mean square error (RMSE) and mean absolute error (MAE) for the 2 models.

Results

Dataset

The comprehensive dataset consisted of 24,949 providers. Table 3 shows the distribution of these providers by specialty and city tier. On one hand, general practice physicians accounted for the largest share of providers by specialty. On the other hand, tier 3 cities accounted for the largest share of providers by city tiers. As shown in Figure 2, the mix of providers by specialty varied across the 3 city tiers (Chi-square*P*P<). The median engagement time for 24,949 providers was 0 min (mean 7.5, IQR 0-1.58), as 69.75% (17,401/24,949) of providers did not consume any content. Moreover, 30.25% (7548/24,949) of providers consumed at least one content piece. The engagement among providers with engagement time greater than 0 min varied significantly, as the median engagement time was 4.9 min (mean 24.8, IQR 2.2-10.1).

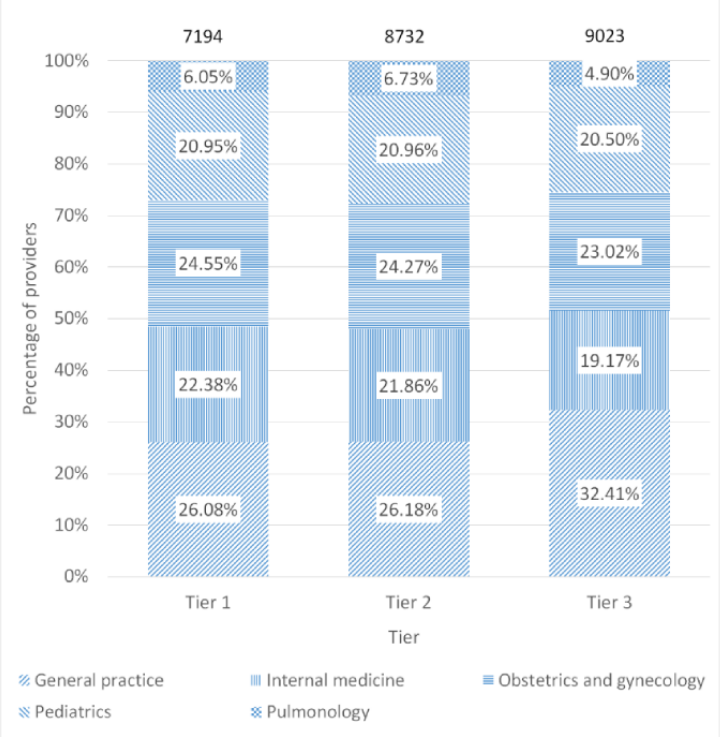
We excluded 69.75% (17,401/24,949) providers who had zero engagement time and further removed 1.36% (339/24,949)

outliers with engagement time greater than the sum of the third quartile and 1.5 times the IQR. As described in the Analysis section, we calculated read proportion, expert proportion, and short proportion for providers. We excluded 2.91% (727/24,949) providers who had undefined proportions (zero divided by zero) for the calculated proportions. Our final dataset contained 6482 providers after removing outliers. The engagement time for these providers, too, varied with median engagement time of 5.2 min (mean 6.6, IQR 2.42-9.83). It also varied by provider specialty and city tier. Among provider specialty, general practice physicians recorded the highest average engagement of 7.2 min, whereas internal medicine physicians recorded the lowest at 6.09 min. Among city tiers, tier 3 providers engaged the highest with the platform, spending an average of 6.8 min, whereas tier 1 providers engaged the lowest at 6.3 min. We conducted one-way analysis of variance tests and found that the engagement time was statistically different across specialty (P<.001) and city tiers (P=.007).

Table 3. Provider participation by city tier and specialty.

City tiers	Specialty, n					
	General practice	Internal medicine	Obstetrics and gynecology	Pediatrics	Pulmonology	Total
Tier 1	1876	1610	1766	1507	435	7194
Tier 2	2286	1909	2119	1830	588	8732
Tier 3	2924	1730	2077	1850	442	9023
Total	7086	5249	5962	5187	1465	24,949

Figure 2. Provider count in the activity dataset by specialty and city tier.



Discovering Content Preferences Among Providers

The average silhouette width was the highest at 0.68 for 14 clusters (Figure 3). However, we chose the number of clusters to be 4, with a marginally lower average silhouette width of 0.66 because of the ease of interpretability of the resulting clusters. Table 4 describes the clusters and their characteristics, which has been explained as follows.

Figure 3. Optimal number of clusters using the average silhouette width method.

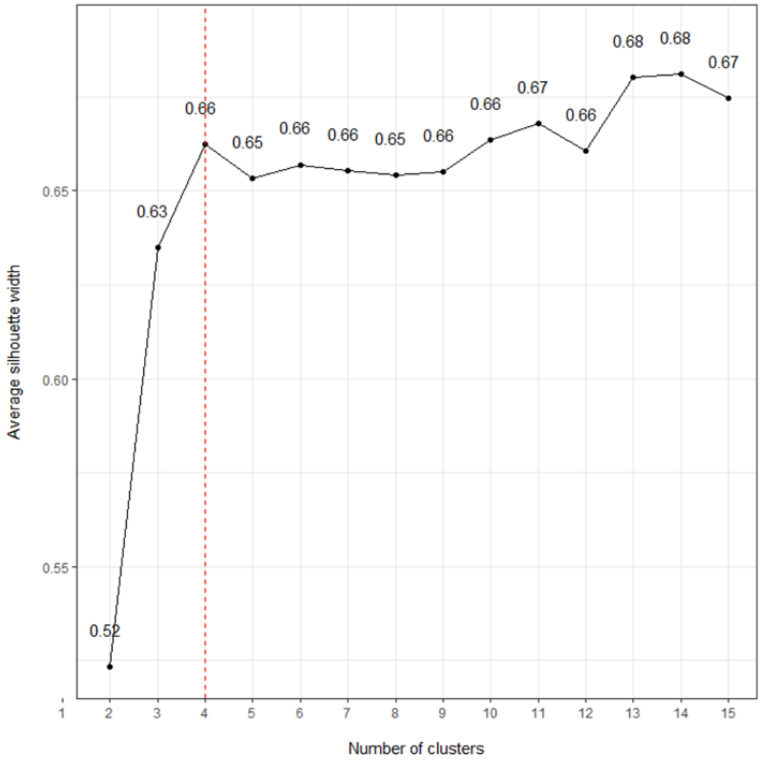


Table 4. Clusters and their characteristics.

Cluster	Providers, n	Engagement time spent reading, %	Engagement time spent on short content, %	Engagement time spent on expert-driven content, %	Average silhouette width
1. Peer-driven microwatchers	2425	1.7	97.3	2.0	0.81
2. Expert-driven microwatchers	772	5.7	98.8	92.7	0.75
3. Peer-driven microreaders	923	50.7	87.5	3.1	0.34
4. Peer-driven long watchers	2362	5.9	20.5	2.3	0.61
Total	6482	10.7	68.1	13.1	0.66

The largest cluster containing 2425 providers was labeled as *peer-driven microwatchers*. It was also the cluster with the highest average silhouette width (0.81), indicating that there was higher homogeneity within the cluster when compared with heterogeneity across other clusters. Providers in this cluster spent 98.27% (8493/8642 min) of their time watching videos. They spent 97.28% (8408/8642 min) of their time engaging with short content. Between video modes, they preferred peer-driven content, as they spent 97.95% (8097/8266 min) of their time on peer-driven content.

The second cluster was labeled as e*xpert-driven microwatchers*, which was the smallest cluster with 772 providers. Providers in this cluster spent 94.27% (2850/3024 min) of their time watching videos and 98.85% (2989/3024 min) of their time engaging with short content. However, in contrast with providers in the first cluster, providers in the second cluster spent 92.67% (2526/2726 min) of their time watching expert-driven videos, indicating that they preferred expert-driven content over peer-driven content.

The third cluster was labeled as *peer-driven microreaders*. This cluster had 923 providers and registered the smallest average silhouette width of 0.34, indicating that the providers in this cluster were less similar among themselves than providers in other clusters. Providers in this cluster spent more time reading (50.65%, 3870/7640 min) than watching (49.35%, 3770/7640 min). They preferred short content and peer-driven videos, as they spent 87.52% (6687/7640 min) on short content and 96.95% (3924/4047 min) of their time on peer-driven videos.

The fourth cluster with 2362 providers was labeled as *peer-driven long watchers*. Providers in this cluster spent 94.06% (22,288/23,696) of their time watching videos and 97.70% (21,535/22,041 min) of their time watching videos that

featured peers. Moreover, they spent 79.51% (18,841/23,696 min) of their time engaging with longer content.

Figures 4 and 5 show the composition of clusters by specialty and city tier, respectively. The proportion of general practice physicians was between 33.5% (259/772) and 44.88% (1060/2362), whereas the proportion of tier 1 providers was between 27.05% (639/2362) and 30.2% (233/772) across all 4 clusters. The Chi-square tests revealed that the city tier mix was similar across the clusters (P=.17), but the specialty mix among the clusters was different (P<.001).

Figure 4. Composition of clusters by specialty.

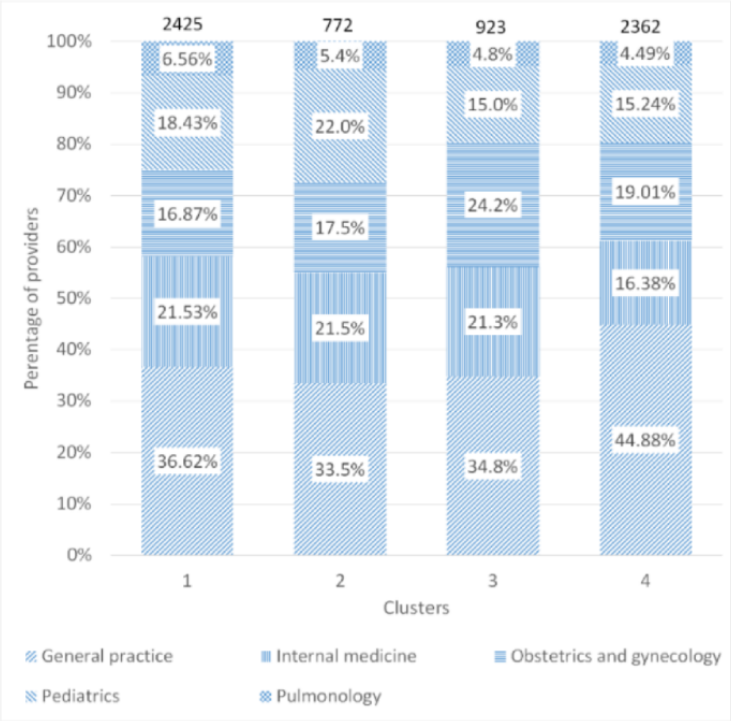
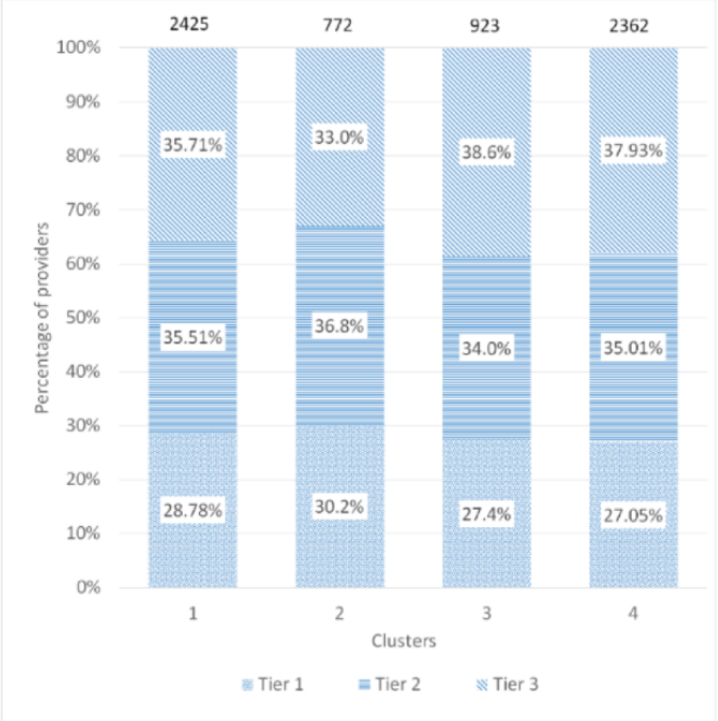


Figure 5. Composition of clusters by city tier.



Predicting Provider Engagement

Table 5 shows the results from the linear regression models that use clusters (model A) and demographic variables (model B) to explain variation in provider engagement. For model A, the R^2 statistic was 0.329, whereas the R^2 statistic for model B was 0.010. In other words, clusters were able to explain a significantly higher proportion of the variation in engagement time when compared with that explained by demographic variables (32.9% vs 1.0%).

Table 5. Regression results for model A (cluster-based model) and model B (demographic variable–based model).

Independent variables	*Dependent variable* (engagement time)	
	Model A- Coefficient (standard error)[a,b]	Model B- Coefficient (standard error)[a,c]
Cluster 2	0.353[d] (0.177)	N/A[e]
Cluster 3	4.713[f] (0.166)	N/A
Cluster 4	6.468[f] (0.124)	N/A
Specialty, internal medicine	N/A	−1.116[f] (0.179)
Specialty, obstetrics and gynecology	N/A	−1.030[f] (0.182)
Specialty, pediatrics	N/A	−0.644[f] (0.187)
Specialty, pulmonology	N/A	−0.902[f] (0.297)
City tier, tier 2	N/A	0.364[d] (0.164)
City tier, tier 3	N/A	0.393[d] (0.162)
Constant	3.564[f] (0.087)	6.934[f] (0.147)

[a]Observations used: 6482
[b]R^2: 0.329; Adjusted R^2: .0329
[c]R^2: 0.010; Adjusted R^2: .0.009
[d]P<.05.
[e]Not applicable.
[f]P<.01.

Table 6 compares the RMSE, R^2 statistic, and MAE from the 2 predictive models based on the 10-fold cross-validation method. In addition to being able to explain a significantly higher proportion of variation in engagement time, we observed that model A also resulted in 17.7% lower RMSE and 22.7% lower MAE than that of model B.

Table 6. Comparison between regression models for engagement time based on 10-fold cross-validation error rates.

Evaluation metrics	Model A based on behavioral variables (clusters)	Model B based on demographic variables (specialty, city tier)	Difference (%; calculated as model B −model A)/model B)
Root mean square error	4.29	5.21	17.7
R^2 statistic	0.33	0.01	−3275.7
Mean absolute error	3.30	4.26	22.7

Discussion

Principal Findings

This study analyzed a mobile phone–based campaign that focused on educating private providers in India on TB diagnostic practices. It was found that there is heterogeneity in provider engagement with ICT-based training interventions. We also found that providers have inherent preferences for the type of instructional content. Content preferences were used to cluster providers, and it was shown that these clusters explain the variation in provider engagement better than the demographic variables such as provider specialty and city tier.

ICT-based public health interventions are often low intensity, have a wide reach, and witness low provider engagement [20]. In our study, providers spent an average of 7.5 min with the yearlong ThinkTB campaign, which is lower than that observed in similar interventions across the world [21]. However, the relatively lower engagement may be because of a significantly wider reach when compared with similar precedents, indicating an engagement vs reach trade-off in provider training interventions. Similar to previous studies, provider engagement with the ThinkTB campaign varied significantly, which could have affected the overall effectiveness of the campaign. The variation in provider engagement is important because there is some evidence that provider engagement and effectiveness of public health interventions are closely related [22]. In particular,

providers who have lower engagement with the intervention often have worse patient outcomes [21,23].

Discovering Content Preferences Among Providers

Our clustering analysis identified groups of providers with homogeneous content preferences despite inherent individual content preferences among providers. Notably, it did not feature all possible combinations of content format, mode, and length. For example, there is no cluster of providers who prefer to watch longer videos featuring experts. This is because, given the number of clusters, the k-means algorithm chooses clusters with the highest average silhouette width across clusters. An increase in the number of clusters would mean poorer interpretability, and thereby, lower feasibility of catering to those additional clusters.

Our clustering analysis confirmed prior research, which shows that health care professionals are known to have individual content preferences [11]. One of the methods used to study instructional content preferences among medical students is the VARK model, which measures preferences for 4 content forms—visual, auditory, reading/writing, and kinesthetic (VARK) [23]. Our study differed from such studies in two ways. First, prior work studied content preferences in an academic context, largely to educate medical students. In contrast, we studied content preferences among practicing providers through a digital campaign geared toward changing their diagnostic behavior. Second, our study identified preferences for unexplored content types that are important for ICT-based interventions to engage with practicing providers in a nonacademic setting. For example, our campaign included 2 of the 4 content forms proposed by the VARK model—visual (videos) and reading/writing (articles). In addition to the content format (video or article), our study also identified preferences for other content types—mode (expert-driven or peer-driven) and length (long or short)—which could provide insights for designing future ICT-based interventions.

Our clustering analysis highlighted two themes that were common for most of the providers. First, providers, on average, preferred shorter content. This is an intuitive outcome, especially for Indian health professionals who experience high burnout rates and work-related stress [24,25]. Given the time constraints, they would presumably prefer to consume shorter content on ICT-based platforms. Second, most providers, on average, preferred watching videos to reading articles. The VARK model suggests that instructional content preferences among medical students vary by study. In particular, there is no clear preference between video and reading/writing content forms among medical students [26-28]. However, our results showed that, on average, providers have a strong preference for video content, as they spent only 10.7% of their time reading articles (Table 4). This difference in results could be because we studied content preferences among practicing providers in a nonacademic setting, unlike studies involving VARK models, which studied content preferences among medical students in an academic setting.

Although the literature is mostly equivocal on the impact of catering to such preferences [29], there is some evidence that customization could result in efficient and effective learning [30]. Within health care settings as well, catering to individual preferences of medical practitioners participating in health care interventions is known to drive the effectiveness of health care interventions [31]. Although content preferences may be individual to every provider, our clustering analysis shows that they could be identified at a group level too.

Predicting Provider Engagement

A comparison of prediction models revealed that clusters based on content preferences predicted engagement time better than demographic variables. We also showed that the composition of clusters varied by specialty but not by city tier, thereby implying that demographic variables may not always be associated with content preferences, and hence, the ability of demographic variables to explain the variation in behavior may be different from that of clusters formed on the basis of content preferences.

At first glance, better prediction of engagement using clusters may seem obvious because clusters were created using a measure of engagement time. However, it is important to note that we used the *proportion of engagement time* spent by providers on various types of content and not the absolute magnitude of the engagement time. This implies that 1 cluster can contain providers with varying levels of engagement (low and high) as long as the proportions of time spent on different content types are similar. Hence, clusters are not guaranteed to provide a better prediction of engagement time by definition.

Recommendations

Using content preferences to engage with providers has important implications for provider-focused training interventions. Provider-focused training interventions often target providers based on demographic variables, such as geographic location (rural, urban, etc), provider specialty (internal medicine, family practice, etc), or clinical setting (university, private practices, etc) [32-34]. Such interventions are unable to recognize and leverage heterogeneity in content preferences at an individual level. Our study provides evidence that suggests moving away from provider demographic information. We propose that training interventions leverage ICT-based platforms to learn individual preferences and deliver customized content based on provider preferences instead of demographic variables.

Learning provider preferences could be operationally challenging, but ICT tools could act as the enabler of the proposed approach. When compared with traditional counterparts, such as physical outreach visits and conferences, ICT-based training interventions are flexible in terms of training timing and sequences and improved access to geographically dispersed providers [35]. Their interactive nature allows dynamic assessment of learners' preferences without which public health interventions resort to using demographic variables for engaging with providers. In addition, ICT-based interventions are adaptable to deliver customized content based on individual content preferences [36]. Therefore, it is feasible to customize instructional content for public health interventions because of ICT-based solutions.

Limitations

Our study has certain limitations. Providers may have left the mobile screen open for a prolonged period without actually engaging with the campaign. Hence, the engagement time recorded on the platform might not reflect the actual time spent by the provider. We partially addressed this issue by removing outliers from our analyses.

Our study offers limited generalizability because of three reasons. First, our clustering analysis and prediction models excluded 69.75% (17,401/24,949) of providers, who had zero engagement time with the campaign. We could not assess the preferences of providers who did not engage with the campaign, which may limit the generalizability of findings to providers who did not engage with the content. A large proportion of unresponsive providers may have also introduced a selection bias. It is possible that providers who did not engage with the campaign did not find ThinkTB relevant or did not engage with the mobile platform at all. However, this did not affect our insights into content preferences, as the heterogeneity in engagement time of providers who had nonzero engagement time allowed us to perform clustering.

Second, the fact that provider clusters in our study were based on individual content preferences limits the generalizability of our results. The content preferences are unique to a provider, which limits the generalizability of our results to other contexts such as another training subject with different content types delivered via a technology platform that is not mobile phone based. Nonetheless, the identification of unique preferences and catering to those preferences is generalizable. On the basis of our results, engaging with providers based on their individual preferences instead of demographic variables could lead to higher engagement with the intervention.

Third, the campaign focused on educating private providers in India on TB diagnostic practices. Providers on the mobile phone–based platform either may not be interested in this topic or may not identify with the platform for this topic. Hence, results from this study may not be generalizable to other clinical contexts other than TB. However, the clinical context of TB in itself is significant in scope. TB is the leading cause of death from a single infectious agent, ranking above HIV/AIDS [37]. Eradication of TB therefore has been a global priority. The World Health Organization (WHO) designed the End TB Strategy in 2014 and subsequently, the United Nations General Assembly included *ending the TB epidemic* as one of the Sustainable Development Goals in 2015 [38]. Moreover, India—the geographical focus of our study—contributes more than a fourth of global TB incidence [37]. The private sector in India is estimated to account for half of the patients with TB in India [39] and is known to be suboptimal in their TB diagnostic and treatment practices [40]. Private sector engagement models in India are being scaled to multiple cities [41], but these physical engagement models are resource intensive [42]. Our campaign was one of the first ICT-based interventions that could provide a more sustainable option of engaging with private providers. Beyond India, too, engaging with private providers is identified as a global priority. WHO asserted that engaging with the private sector could account for 3.6 million missing TB cases globally and proposed the adoption of the Public-Private Mix model to improve TB detection and treatment [43]. Therefore, our study has implications for interventions aimed at engaging with private providers for TB care across the globe.

Conclusions

Our study shows that providers participating in a mobile phone–based digital campaign have inherent preferences for the instructional content. It also shows that targeting providers by catering to individual provider content preferences could result in a higher provider engagement when compared with targeting them based on demographic variables. A higher provider engagement could maximize provider learning and improve the effectiveness of public health interventions. ICT allows us to cater to individual content preferences and could be leveraged to design provider-centric health interventions.

Authors' Contributions

MS and ND designed and executed the ThinkTB campaign in collaboration with the third-party mobile phone–based platform under the supervision of SD and AP. SD and HT conceptualized the study. HT conducted the analysis and wrote the initial draft. SD edited and revised subsequent versions of the draft. All authors critically reviewed the draft and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008 Aug 25;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
2. de Angelis G, Davies B, King J, McEwan J, Cavallo S, Loew L, et al. Information and communication technologies for the dissemination of clinical practice guidelines to health professionals: a systematic review. JMIR Med Educ 2016 Nov 30;2(2):e16 [FREE Full text] [doi: 10.2196/mededu.6288] [Medline: 27903488]
3. George PP, Papachristou N, Belisario JM, Wang W, Wark PA, Cotic Z, et al. Online eLearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. J Glob Health 2014 Jun;4(1):010406 [FREE Full text] [doi: 10.7189/jogh.04.010406] [Medline: 24976965]

4. Glasgow RE, Emmons KM. How can we increase translation of research into practice? Types of evidence needed. Annu Rev Public Health 2007;28:413-433. [doi: 10.1146/annurev.publhealth.28.021406.144145] [Medline: 17150029]
5. George PP, Zhabenko O, Kyaw BM, Antoniou P, Posadzki P, Saxena N, et al. Online Digital Education for Postregistration Training of Medical Doctors: Systematic Review by the Digital Health Education Collaboration. J Med Internet Res 2019 Feb 25;21(2):e13269 [FREE Full text] [doi: 10.2196/13269] [Medline: 30801252]
6. Bennett GG, Glasgow RE. The delivery of public health interventions via the internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
7. Epstein JN, Langberg JM, Lichtenstein PK, Kolb R, Altaye M, Simon JO. Use of an internet portal to improve community-based pediatric ADHD care: a cluster randomized trial. Pediatrics 2011 Nov;128(5):e1201-e1208 [FREE Full text] [doi: 10.1542/peds.2011-0872] [Medline: 22007005]
8. Liu Q, Peng W, Zhang F, Hu R, Li Y, Yan W. The effectiveness of blended learning in health professions: systematic review and meta-analysis. J Med Internet Res 2016 Jan 4;18(1):e2 [FREE Full text] [doi: 10.2196/jmir.4807] [Medline: 26729058]
9. Ally M. Foundations of educational theory for online learning. In: Anderson T, editor. The Theory and Practice of Online Learning. Athabasca, Alberta: Athabasca University Press; 2004:15-44.
10. Rovai AP. Building sense of community at a distance. Int Rev Res Open Dist Learn 2002 Apr;3(1):1-16. [doi: 10.19173/irrodl.v3i1.79]
11. Andrusyszyn MA, Cragg CE, Humbert J. Nurse practitioner preferences for distance education methods related to learning style, course content, and achievement. J Nurs Educ 2001 Apr;40(4):163-170. [Medline: 11324833]
12. Rohrer D, Pashler H. Learning styles: where's the evidence? Med Educ 2012 Jul;46(7):634-635. [doi: 10.1111/j.1365-2923.2012.04273.x] [Medline: 22691144]
13. McMahon GT. What do I need to learn today?--The evolution of CME. N Engl J Med 2016 Apr 14;374(15):1403-1406. [doi: 10.1056/NEJMp1515202] [Medline: 27074064]
14. Mishra G, Mulani J. Tuberculosis prescription practices in private and public sector in India. Natl J Integr Res Med 2013;4(2):71-78 [FREE Full text]
15. Ministry of Labour and Employment. Ministry of Labour & Employment | GoI. 2017. The Gazette of India: Extraordinary: Annexure S.O. 186(E) URL: https://labour.gov.in/sites/default/files/MW%20Final%20%281%29_0.pdf [accessed 2020-01-14]
16. The R Project for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing URL: http://www.R-project.org/ [accessed 2020-01-14]
17. Hartigan JA, Wong MA. Algorithm AS 136: A k-means clustering algorithm. J R Stat Soc Ser C Appl Stat 1979;28(1):100-108. [doi: 10.2307/2346830]
18. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20:53-65. [doi: 10.1016/0377-0427(87)90125-7]
19. Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In: Proceedings of the 14th international joint conference on Artificial intelligence. 1995 Presented at: IJCAI'95; August 20-25, 1995; Montreal, Quebec p. 1137-1143.
20. Glasgow RE, Klesges LM, Dzewaltowski DA, Estabrooks PA, Vogt TM. Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. Health Educ Res 2006 Oct;21(5):688-694. [doi: 10.1093/her/cyl081] [Medline: 16945984]
21. Estrada CA, Safford MM, Salanitro AH, Houston TK, Curry W, Williams JH, et al. A web-based diabetes intervention for physician: a cluster-randomized effectiveness trial. Int J Qual Health Care 2011 Dec;23(6):682-689 [FREE Full text] [doi: 10.1093/intqhc/mzr053] [Medline: 21831967]
22. Billue KL, Safford MM, Salanitro AH, Houston TK, Curry W, Kim Y, et al. Medication intensification in diabetes in rural primary care: a cluster-randomised effectiveness trial. BMJ Open 2012;2(5):pii: e000959 [FREE Full text] [doi: 10.1136/bmjopen-2012-000959] [Medline: 22991217]
23. Crenshaw K, Curry W, Salanitro AH, Safford MM, Houston TK, Allison JJ, et al. Is physician engagement with Web-based CME associated with patients' baseline hemoglobin A1c levels? The Rural Diabetes Online Care study. Acad Med 2010 Sep;85(9):1511-1517 [FREE Full text] [doi: 10.1097/ACM.0b013e3181eac036] [Medline: 20736679]
24. Jugale PV, Mallaiah P, Krishnamurthy A, Sangha R. Burnout and work engagement among dental practitioners in Bangalore city: A cross-sectional study. J Clin Diagn Res 2016 Feb;10(2):ZC63-ZC67 [FREE Full text] [doi: 10.7860/JCDR/2016/16009.7312] [Medline: 27042589]
25. Ratnakaran B, Prabhakaran A, Karunakaran V. Prevalence of burnout and its correlates among residents in a tertiary medical center in Kerala, India: A cross-sectional study. J Postgrad Med 2016;62(3):157-161 [FREE Full text] [doi: 10.4103/0022-3859.184274] [Medline: 27320952]
26. Prithishkumar IJ, Michael SA. Understanding your student: using the VARK model. J Postgrad Med 2014;60(2):183-186 [FREE Full text] [doi: 10.4103/0022-3859.132337] [Medline: 24823519]
27. James S, D'Amore A, Thomas T. Learning preferences of first year nursing and midwifery students: utilising VARK. Nurse Educ Today 2011 May;31(4):417-423. [doi: 10.1016/j.nedt.2010.08.008] [Medline: 20826044]

28. Alkhasawneh IM, Mrayyan MT, Docherty C, Alashram S, Yousef HY. Problem-based learning (PBL): assessing students' learning preferences using VARK. Nurse Educ Today 2008 Jul;28(5):572-579. [doi: 10.1016/j.nedt.2007.09.012] [Medline: 17983691]
29. Pashler H, McDaniel M, Rohrer D, Bjork R. Learning styles: concepts and evidence. Psychol Sci Public Interest 2008 Dec;9(3):105-119. [doi: 10.1111/j.1539-6053.2009.01038.x] [Medline: 26162104]
30. Carver Jr CA, Howard R, Lavelle E. Enhancing Student Learning by Incorporating Learning Styles Into Adaptive Hypermedia. In: Proceedings of World Conference on Educational Multimedia and Hypermedia & World Conference on Educational Telecommunications.: Association for the Advancement of Computing in Education (AACE); 1996 Presented at: EdMedia+ Innovate Learning'96; June 17-22, 1996; Boston, USA p. 28-32.
31. Rothman AJ, Baldwin AS. A person x intervention strategy approach to understanding health behavior. In: Deaux K, Snyder M, editors. The Oxford Handbook of Personality and Social Psychology. Oxford, United Kingdom: Oxford University Press; 2012.
32. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. J Am Med Assoc 1995 Sep 6;274(9):700-705. [doi: 10.1001/jama.274.9.700] [Medline: 7650822]
33. Beach MC, Gary TL, Price EG, Robinson K, Gozu A, Palacio A, et al. Improving health care quality for racial/ethnic minorities: a systematic review of the best evidence regarding provider and organization interventions. BMC Public Health 2006 Apr 24;6:104 [FREE Full text] [doi: 10.1186/1471-2458-6-104] [Medline: 16635262]
34. Szilagyi PG, Schaffer S, Shone L, Barth R, Humiston SG, Sandler M, et al. Reducing geographic, racial, and ethnic disparities in childhood immunization rates by using reminder/recall interventions in urban primary care practices. Pediatrics 2002 Nov;110(5):e58. [doi: 10.1542/peds.110.5.e58] [Medline: 12415064]
35. Fordis M, King JE, Ballantyne CM, Jones PH, Schneider KH, Spann SJ, et al. Comparison of the instructional efficacy of Internet-based CME with live interactive CME workshops: a randomized controlled trial. J Am Med Assoc 2005 Sep 7;294(9):1043-1051. [doi: 10.1001/jama.294.9.1043] [Medline: 16145024]
36. Horton W. Designing Web-Based Training: How to Teach Anyone Anything Anywhere Anytime. New York, NY: Wiley; 2000.
37. World Health Organization.: WHO; 2017. Global Tuberculosis Report 2017 URL: https://www.who.int/tb/publications/global_report/gtbr2017_main_text.pdf [accessed 2020-01-14]
38. World Health Organization. 2015. Implementing the End TB Strategy: The Essentials URL: https://www.who.int/tb/publications/2015/end_tb_essential.pdf [accessed 2020-01-14]
39. Hazarika I. Role of private sector in providing tuberculosis care: evidence from a population-based survey in India. J Glob Infect Dis 2011 Jan;3(1):19-24 [FREE Full text] [doi: 10.4103/0974-777X.77291] [Medline: 21572604]
40. Das J, Kwan A, Daniels B, Satyanarayana S, Subbaraman R, Bergkvist S, et al. Use of standardised patients to assess quality of tuberculosis care: a pilot, cross-sectional study. Lancet Infect Dis 2015 Nov;15(11):1305-1313 [FREE Full text] [doi: 10.1016/S1473-3099(15)00077-8] [Medline: 26268690]
41. finddx.: Center for Health Research and Innovation, FIND, Clinton Health Access Initiative; 2018. Joint Effort for Elimination of Tuberculosis (JEET) URL: https://www.finddx.org/wp-content/uploads/2019/05/JEET-Annual-Report-2018.pdf [accessed 2020-01-14]
42. Deo S, Jindal P, Gupta D, Khaparde S, Rade K, Sachdeva KS, et al. What would it cost to scale-up private sector engagement efforts for tuberculosis care? Evidence from three pilot programs in India. PLoS One 2019;14(6):e0214928 [FREE Full text] [doi: 10.1371/journal.pone.0214928] [Medline: 31166942]
43. World Health Organization.: WHO; 2018. Public-Private Mix for TB Prevention and Care URL: https://www.who.int/tb/publications/2018/PPMRoadmap.pdf [accessed 2020-01-14]

Abbreviations

ICT: information and communications technology
MAE: mean absolute error
RMSE: root mean square error
TB: tuberculosis
VARK: visual, auditory, reading/writing, and kinesthetic
WHO: World Health Organization

Edited by G Eysenbach; submitted 27.08.19; peer-reviewed by L Tudor Car, Y Ariza-Araujo, P Antoniou; comments to author 28.10.19; revised version received 29.11.19; accepted 16.12.19; published 03.03.20.

Please cite as:
Tyagi H, Sabharwal M, Dixit N, Pal A, Deo S
Leveraging Providers' Preferences to Customize Instructional Content in Information and Communications Technology–Based Training Interventions: Retrospective Analysis of a Mobile Phone–Based Intervention in India
JMIR Mhealth Uhealth 2020;8(3):e15998
URL: https://mhealth.jmir.org/2020/3/e15998
doi:10.2196/15998
PMID:32130191

Original Paper

Efficacy and Safety of an mHealth App and Wearable Device in Physical Performance for Patients With Hepatocellular Carcinoma: Development and Usability Study

Yoon Kim[1], MD, MMSc; Jinserk Seo[2], BPTh; So-Yeon An[3], MPT; Dong Hyun Sinn[4], MD, PhD; Ji Hye Hwang[1], MD, PhD

[1]Department of Physical and Rehabilitation Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[2]Department of Rehabilitation Medicine, Incheon St. Mary's Hospital, The Catholic University of Korea, Incheon, Republic of Korea

[3]Department of Health Science, Korea University Graduate School, Korea University, Seoul, Republic of Korea

[4]Department of Internal Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

Corresponding Author:
Ji Hye Hwang, MD, PhD
Department of Physical and Rehabilitation Medicine
Samsung Medical Center
Sungkyunkwan University School of Medicine
Irwon-Ro 81
Gangnam-Gu
Seoul, 06351
Republic of Korea
Phone: 82 2 3410 2818
Email: jhlee.hwang@samsung.com

Abstract

Background: Exercise is predicted to have a positive effect among hepatocellular carcinoma (HCC) patients. However, these patients are hesitant to start and build up an exercise program for one major reason: the vague fear of developing hepatic decompensation, a potentially fatal condition that can lead to death. Integrating mobile health (mHealth) with individualized exercise programs could be a possible option for promoting physical capacity among HCC patients.

Objective: The aim of this study was to evaluate the efficacy and safety of rehabilitation exercises, which have been individually prescribed via an mHealth app, on physical fitness, body composition, biochemical profile, and quality of life among HCC patients.

Methods: A total of 37 HCC patients were enrolled in a 12-week course with an mHealth app program targeted to HCC patients. The wearable wristband device Neofit (Partron Co) was provided to participants, and recorded daily physical data, such as the number of steps, calorie expenditure, exercise time, and heart rate. Each participant was given an individualized rehabilitation exercise program that was prescribed and adjusted at the 6-week midintervention period based on the assessment results. At baseline, 6-week, and 12-week sessions, participants' physical fitness levels (ie, 6-minute walk test, grip strength test, and 30-second chair stand test) were measured. Physical activity levels, as measured by the International Physical Activity Questionnaire-Short Form (IPAQ-SF); body composition (ie, body mass index, body fat percentage, and muscle mass); biochemical profiles; and quality of life, as measured by the European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire C30, were assessed at baseline and at the end point. At the 6-week midpoint, exercise intensity was individually adjusted.

Results: Of the 37 patients, 31 (84%) completed the 12-week intervention. Grip strength improved significantly after 12 weeks of the intervention. The 30-second chair stand test and the 6-minute walk test showed significant improvement from 0 to 6 weeks, from 0 to 12 weeks, and from 6 to 12 weeks. Muscle mass and the IPAQ-SF score increased significantly after 12 weeks of the intervention without biochemical deterioration.

Conclusions: Following 12 weeks of mHealth care, including an individually prescribed rehabilitation exercise program, we saw significant improvements in physical fitness, body composition, and physical activity without any complication or biochemical deterioration among compensated HCC patients who had completed therapy.

(JMIR Mhealth Uhealth 2020;8(3):e14435) doi:10.2196/14435

KEYWORDS

mHealth; hepatocellular carcinoma; rehabilitation; exercise; physical fitness; physical activity

Introduction

Physical activity has been proven to have a positive influence, both biologically and functionally, in patients with cancer, such as breast or prostate cancer [1-3]. Recent studies indicate that even patients undergoing acute cancer treatments can benefit from individualized exercise programs [4]. However, little is known about the role of exercise on hepatocellular carcinoma (HCC), which is the fifth- and seventh-most common cancer worldwide in men and women, respectively [5].

Among the few studies regarding the effect of exercise in HCC, one study reported that continuous regular exercise improved physical ability without deteriorating liver function in HCC patients with chronic liver disease (CLD) [6]. An experimental study using a rat model found that regular physical activity reduced the risk of the primary development of HCC [7]. Although the positive effect of exercise is quite predictable in HCC patients, these patients are hesitant to start and build up an exercise program for one major reason: the vague fear of developing hepatic decompensation. Hepatic decompensation is a potentially fatal condition that can be encountered by patients with HCC, which includes hepatic encephalopathy, esophageal varices, or ascites. Therefore, in order to safely boost exercise capacity, a delicate supporting system is needed to monitor physical activity and alert such vulnerable patients before hepatic decompensation occurs.

Integrating mobile health (mHealth) with individualized exercise programs could be a possible option for promoting physical capacity of HCC patients. Following the explosive increase in smartphone penetration globally, mHealth has been spotlighted as a novel technology for promoting exercise programs, not only among healthy populations but also among patients with various diseases, such as diabetes, heart disease, and cancer [8-10]. Several recent studies have reported the effectiveness of mHealth in patients with solid cancers and survivors of breast cancer and colorectal cancers [4,11]. mHealth promotes physical activity among cancer patients by motivating them to exercise and by providing real-time feedback [12,13]. Because mHealth can provide large amounts of information and even lower the barrier of communication with health care providers, it is an ideal tool for helping HCC patients to exercise safely outside the hospital.

In this study, we evaluated the efficacy and safety of exercises that were individually prescribed to compensated HCC patients after anticancer therapy. We analyzed changes in physical fitness, body composition, biochemical profile, and quality of life (QoL) after 12 weeks of an mHealth exercise intervention.

Methods

Participants

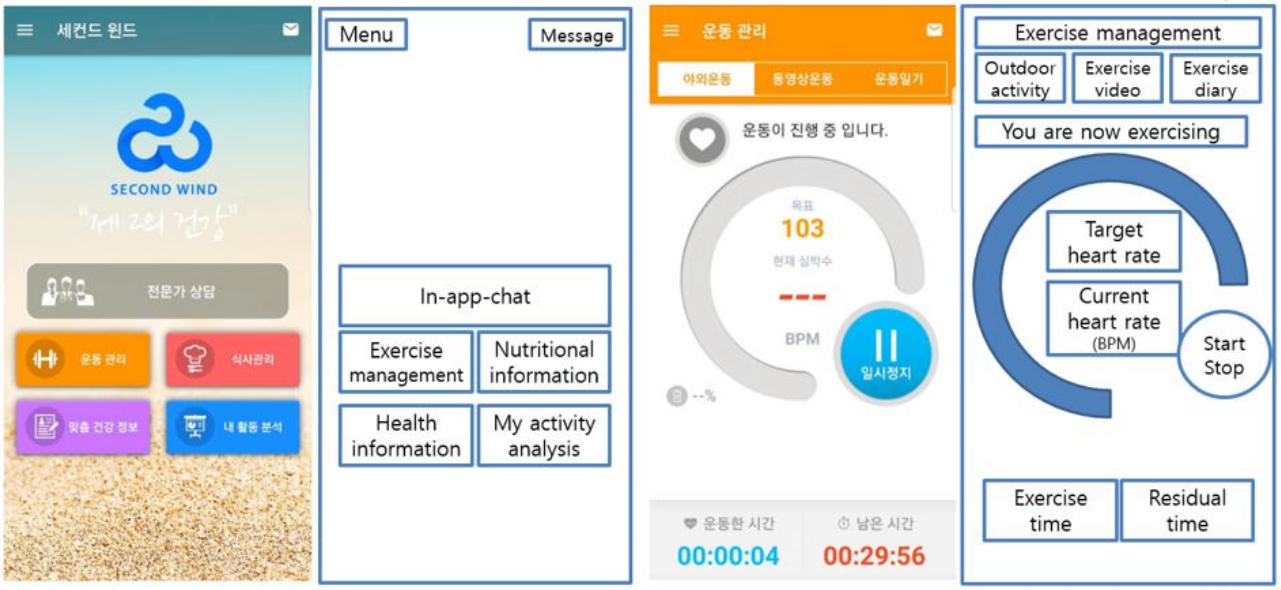
HCC patients who visited the outpatient cancer rehabilitation clinic of a tertiary hospital from November 2017 to February 2018 were prospectively enrolled in this study. The inclusion criteria were as follows: HCC patients aged 19-69 years, patients at stage I or II of the modified Union for International Cancer Control (mUICC) staging system, patients with a Child-Pugh class A or B score, patients who could walk independently for 30 minutes, and patients who had a mobile phone. Exclusion criteria were as follows: patients who required exercise restrictions for severe cardiopulmonary or renal disease, patients with musculoskeletal or neurological deficits, patients with cognitive impairments that interfered with mobile phone utilization, or patients who were unable to give written consent.

App Development

Based on our previous experience of, and knowledge about, developing mHealth apps for specific cancer patients, we recruited app engineers and five health care professionals from a comprehensive cancer center [4]. Health care professionals were recruited to help in the development of a comprehensive mHealth care system tailored to HCC patients, in order to provide health information, self-monitoring, and connections with health care professionals. The app included the following features: *exercise management*, *nutritional information*, *health information*, *my activity analysis*, and *in-app chat* service (see Figure 1).

The exercise management feature provided daily, personalized aerobic and anaerobic exercises based on clinical evidence. The health information content changed daily, including general health information about HCC, medication, adverse effects of anticancer therapy, and nutrition. Real-time communication with a medical professional was available through the in-app chat service. The wearable Internet of Things (IoT) device, which connected with the mobile app, gathered real-time physical data to monitor both physical activity and the health of the participants. A clinical evidence-based care system with the above functions was implemented through the mHealth app.

Figure 1. App screenshots and configuration. Screenshots of the main page (far left) and exercise management page (middle right) are shown, along with respective diagrams of explanations to the right of each screenshot.



Intervention

mHealth Care App

All participants received a 12-week course of an individually prescribed exercise program through the mHealth app using their own mobile phones and an interconnected wearable IoT device worn on their wrists. The participants were told to download the mobile app to their mobile phones. The mobile app consisted of general medical information on HCC, disease-specific exercise care, nutritional information, and a real-time chat service that directly connected participants to the study coordinator. The wearable IoT device—Neofit (Partron Co)—which we provided to the participants, recorded daily physical data, such as the number of steps, calorie expenditure, exercise time, and heart rate. Each participant was given an individualized rehabilitation exercise program that was prescribed and adjusted at the 6-week midintervention period based on the assessment results.

Individually Prescribed Rehabilitation Exercises

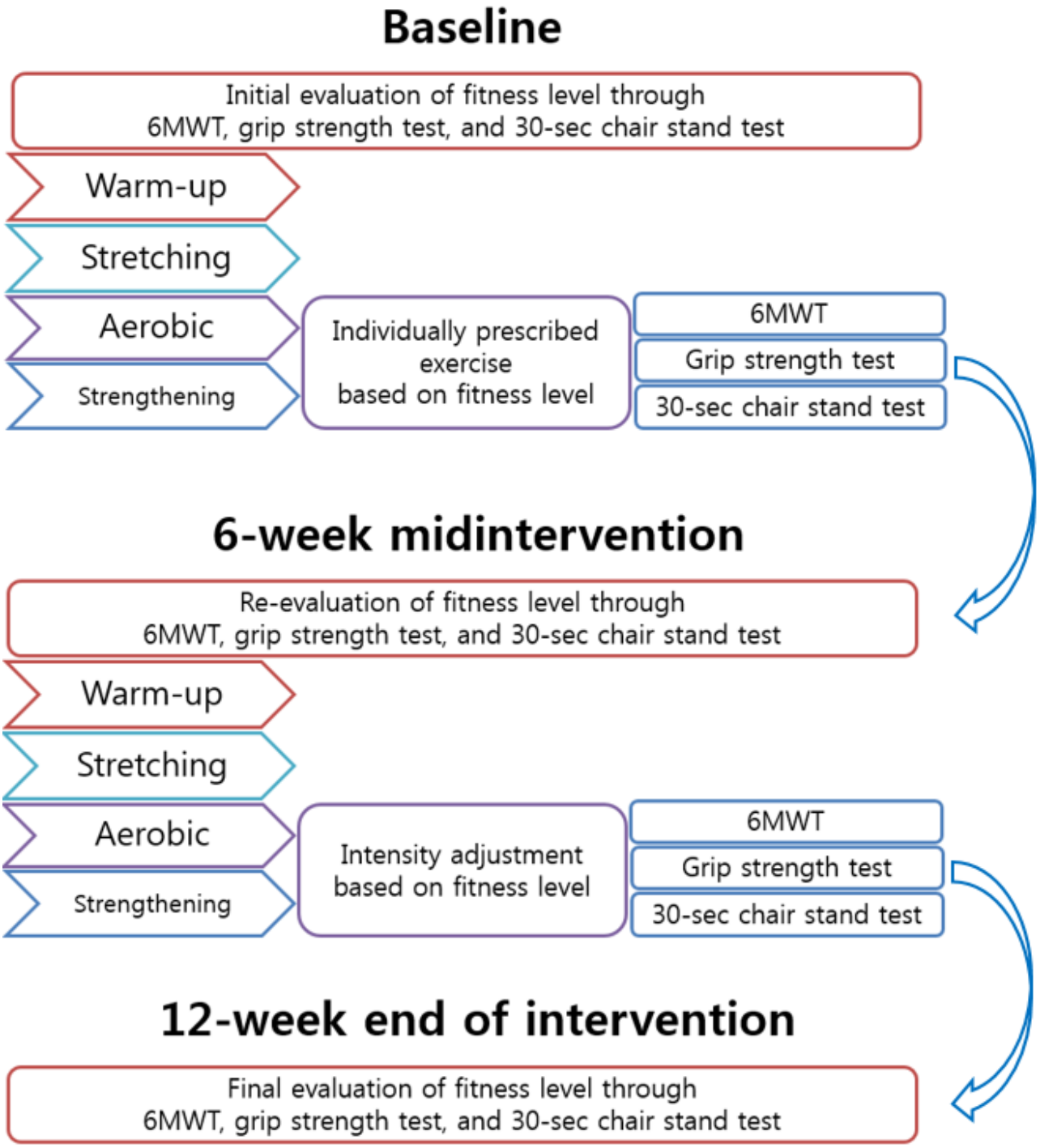
The patients were recommended to participate in individually tailored regular rehabilitation exercises. Video clips were provided by the app and were composed of warm-up, stretching, aerobic, and muscle-strengthening exercises for the upper and lower extremities. Participants were asked to watch the video clips and perform the exercises daily. During the baseline assessment, participants received instructions for standardized warm-up and stretching exercises and individualized aerobic and muscle-strengthening exercises. The aerobic and muscle-strengthening exercise prescriptions were modified based on each participant's altered physical fitness levels measured at the 6-week midpoint session (see Figure 2).

The intensity (ie, light walking, light running, mountain climbing, and cycling) and target heart rate for the aerobic exercise were set from the results of the 6-minute walk test (6MWT). This result was compared with an individualized reference value calculated using each participant's age, height, and weight [14]. If the 6MWT result was higher than the individualized reference value, the recommended exercise intensity and target heart rate were increased.

All the major muscle groups of the upper extremities, lower extremities, and trunk were included in the muscle-strengthening exercise program. Three steps of resistance exercises (ie, maximum, moderate, and minimum resistance) were provided. The results of the 30-second chair stand test and grip strength test were compared with reference values based on the healthy, normal, Korean population from the Korea Sports Promotion Foundation. Appropriate intensity levels for the strengthening exercises were determined according to those results.

After 6 weeks of participation, fitness levels (ie, 6MWT, grip strength test, and 30-second chair stand test) were assessed in all participants. According to the results of this midpoint assessment, the intensity of the aerobic exercise and level of strengthening exercises were adjusted. The target intensity and target heart rate of the aerobic exercise were altered according to the results of the 6MWT. We increased the recommended exercise level if the participant's 6MWT distance increased. The level of the resistance exercises was altered according to the results of the grip strength test and the 30-second chair stand test. Participants' conditions, as a result of their exercise programs, were communicated through the app's real-time chat service. Participant compliance with the exercise program was checked by calculating the total aerobic exercise time completed and the number of resistance exercise video clips watched.

Figure 2. Procedure for adjusting participants' individually prescribed rehabilitation exercise programs in the mHealth app. 6MWT: 6-minute walk test.



Measures

At baseline and at the end point of the study, we collected demographic characteristics and related medical data (ie, blood tests and body composition analyses); in addition, participants completed questionnaires for physical activity and QoL. The outcome measures for physical fitness level were performed at baseline, midintervention (ie, 6 weeks), and at the intervention end point (ie, 12 weeks).

Blood tests, completed at baseline and at the study end point, consisted of a complete blood count (ie, white blood cell count, hemoglobin, and platelets), chemistry panel (ie, albumin, cholesterol, total bilirubin, creatinine, and osteocalcin), liver function tests (ie, aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase, and gamma glutamyl transferase), and a coagulation study (ie, prothrombin time [international normalized ratio]).

Data for the body composition analyses were collected using a bioelectric impedance device—Inbody 720 (Biospace). Muscle mass and body fat percentages were obtained. Body mass index (BMI) was calculated as body weight/height (kg/m^2).

The International Physical Activity Questionnaire-Short Form (IPAQ-SF) was used to measure participants' physical activity

[15]. This questionnaire contains nine questions about the minutes per day or days per week spent doing activities of vigorous and moderate intensity and about time spent walking or sitting during the past 7 days. Using this questionnaire, we calculated the total number of metabolic equivalents (METs) per week. The calculated physical activity was classified into three groups according to the IPAQ-SF scoring system: inactive, minimally active, and highly active. The inactive group consumed less than 600 METs of physical activity, the minimally active group consumed more than 600 METs but less than 3000 METs, and the highly active group consumed a minimum of 3000 METs.

QoL related to general health was assessed by the European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire C30 (EORTC-QLQ-C30) [16]. This questionnaire contains 30 items regarding general health status, five functional scales (ie, physical, role, cognitive, emotional, and social functioning), three symptom scales (ie, fatigue, pain, and nausea or vomiting), and six single-item scales (ie, dyspnea, appetite loss, constipation, diarrhea, financial difficulties, and insomnia). Each scale includes a different set of items, which are calculated using specific coding procedures. Higher scores for the general health status and the functional scales imply positive results, whereas the symptom and the single-item scales are interpreted inversely [17].

Physical fitness was measured with the grip strength test, the 30-second chair stand test, and the 6MWT. A hand-held dynamometer—SH 5001 (Saehan Corp)—was used to assess upper-extremity muscle strength. In an upright posture with slightly abducted arm and slightly flexed elbow, the participants were instructed to hold the dynamometer with the arm and wrist both in a neutral position. After holding it for 3 seconds, maximal power was measured. This grip strength test was repeated three times and the average power was recorded [18]. Lower-limb strength was assessed by the 30-second chair stand test, which counts the maximum number of times a participant can stand up from a chair in 30 seconds. At the start, each participant was seated straight up in a chair without leaning on the backrest and with both arms folded across the chest. For 30 seconds, they repeated complete stand-up and sit-down motions as quickly as possible, and the total number of complete standing motions was counted [19]. To measure the level of cardiopulmonary endurance, the 6MWT was conducted in a 15.2-meter hallway. The total distance walked at maximal speed for 6 minutes was recorded. The physical fitness data were compared with age-specific, normal, Korean fitness values from the *National Fitness 100* project from the Korean government [20].

Statistical Analysis

SPSS Statistics for Windows, version 24.0 (IBM Corp), was used for statistical analyses, and statistical significance was established as $P<.05$. General and clinical subject characteristics were analyzed using descriptive statistics. One-way, repeated-measures analysis of variance (ANOVA) was used to analyze changes in physical fitness over time. Because body composition, physical activity levels, and biochemical profiles were checked at baseline and after 12 weeks of the intervention, the paired *t* test or the Wilcoxon signed-rank test was used to determine the effects of the exercise intervention.

Ethics Approval and Consent to Participate

All decisions regarding this study were approved by the Institutional Review Board of Samsung Medical Center after a complete review of clinical trial protocols (approval number: 2017-06-050). All participants provided written informed consent.

Results

Demographics and Clinical Characteristics

A total of 37 patients diagnosed with HCC were enrolled in this study. Among the 37 patients, 31 of them—26 (84%) males and 5 (16%) females—completed the 12-week intervention using the mHealth app program on a mobile phone with an interconnected wearable device. The 6 patients (6/37, 16%) who did not use the mHealth app and did not respond to our communications regarding the 6-week midintervention evaluation were removed from the study. The demographic and clinical characteristics of the study participants are presented in Table 1. Their mean age was 56.7 years (SD 7.7) and 84% (26/31) of them were male. Participants' mean BMI was 25.39 kg/m^2 (SD 3.00) and mean muscle mass was 28.98 kg (SD 5.20). Participants' initial mean body fat percentage was 26.29% (SD 8.01), which is in the upper-normal range (normal range: 18.0%-28.0%). As for underlying CLD, 24 out of 31 (77%) patients had liver cirrhosis and 7 (23%) had chronic hepatitis. In total, 74% (23/31) of our patients had been diagnosed with HCC for more than 1 year, and 4 patients out of 31 (13%) were long-term patients of more than 5 years. In total, 19% (6/31) of patients had previously experienced hepatic decompensation, mainly variceal hemorrhage. The most common therapeutic method was surgery (17/31, 55%), followed by combination treatments (11/31, 35%) and locoregional therapies (ie, radiofrequency ablation or transarterial chemoembolization) (3/31, 10%).

Table 1. Demographic and clinical characteristics of participants, (N=31).

Characteristic	Value
Age (years), mean (SD)	56.7 (7.7)
Gender, n (%)	
Male	26 (84)
Female	5 (16)
Height (cm), mean (SD)	166.9 (8.2)
Weight (kg), mean (SD)	71.0 (11.3)
Body mass index (kg/m^2), mean (SD)	25.39 (3.00)
Body fat (%), mean (SD)	26.29 (8.01)
Muscle mass (kg), mean (SD)	28.98 (5.20)
Underlying chronic liver disease, n (%)	
Liver cirrhosis	24 (77)
Chronic viral hepatitis	5 (16)
Nonalcoholic fatty liver disease	1 (3)
Alcoholic liver disease	1 (3)
Diagnosis date, n (%)	
0-6 months ago	5 (16)
6 months-1 year ago	3 (10)
1-3 years ago	11 (35)
3-5 years ago	8 (26)
More than 5 years ago	4 (13)
Comorbidity, n (%)	
Diabetes mellitus	7 (23)
Hypertension	10 (32)
Dyslipidemia	1 (3)
Cardiopulmonary disease	1 (3)
Previous experience of hepatic decompensation, n (%)	
None	25 (81)
More than once	6 (19)
Treatment, n (%)	
Locoregional therapies (ie, transarterial chemoembolization or radiofrequency ablation)	3 (10)
Surgery	17 (55)
Combination treatment	11 (35)

Physical Fitness Measures and Compliance Rate

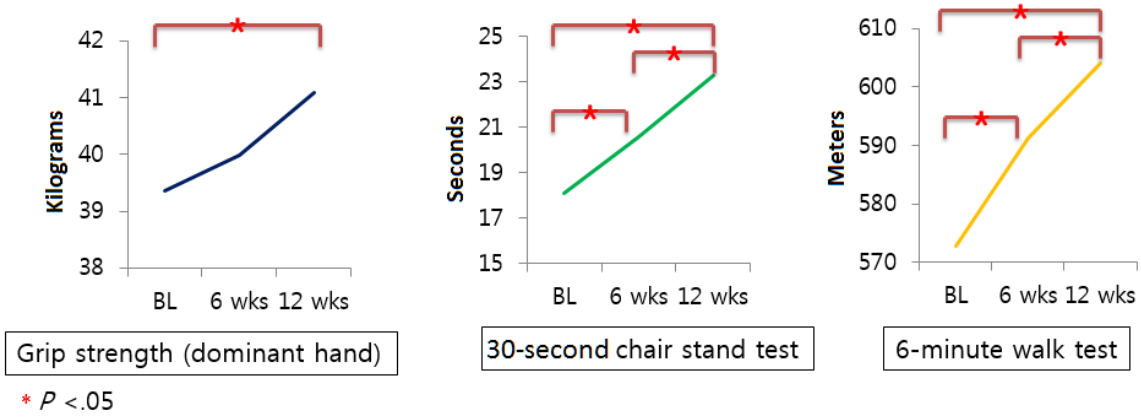
Table 2 and Figure 3 present the serial changes in the objective physical fitness measures at baseline and after using the mHealth app with individually prescribed rehabilitation exercises for 6 and 12 weeks. From baseline to final measurement, grip strength was graded as poor according to the Korean national fitness normal value for 55-59-year-old males. However, compared with baseline, grip strength did improve significantly after 12 weeks of the intervention (P=.02). The 30-second chair stand test also showed significant improvement after 6 and 12 weeks, compared with baseline (P<.001 and P<.001, respectively), and from 6 to 12 weeks (P<.001). Likewise, the 6MWT showed significant improvement after 6 and 12 weeks, compared with baseline (P<.001 and P<.001, respectively), and from 6 to 12 weeks (P=.01).

For aerobic exercise, the stiffest decline in compliance took place at 6-7 weeks, whereas for the strengthening exercises, the most rapid decline was at 4-5 weeks (δ=9.68% and δ=11.29%, respectively).

Table 2. Changes in physical fitness measures.

Measure	Baseline, mean (SD)	6 weeks, mean (SD)	12 weeks, mean (SD)	Baseline vs 6 weeks, *P* value	6 weeks vs 12 weeks, *P* value	Baseline vs 12 weeks, *P* value
Grip strength (kg)	39.35 (9.98)	39.98 (9.69)	41.10 (10.52)	.69	.31	.02
30-second chair stand test (seconds)	18.10 (3.29)	20.55 (3.03)	23.26 (3.79)	<.001	<.001	<.001
6-minute walk test (meters)	572.90 (49.15)	591.19 (49.61)	604.07 (51.59)	<.001	.01	<.001

Figure 3. Changes in physical fitness measures. BL: baseline; wks: weeks.



Body Composition and Self-Reported Physical Activity

As shown in Table 3 and Figure 4, muscle mass increased significantly after 12 weeks of prescribed exercise (*P*=.03). The change in BMI was insignificant, and body fat percentage declined but without statistical significance (*P*=.08 and *P*=.51, respectively). At baseline, the mean, weekly, physical activity score reported on the IPAQ-SF was 2031.95 (SD 2236.60), placing participants in the minimally active group. After completing 12 weeks of the exercise intervention, the mean IPAQ-SF score increased to 3479.71 (SD 2640.08), which places participants in the highly active group. The increase in the IPAQ-SF score was significant after 12 weeks of the intervention (*P*=.01) (see Table 3 and Figure 4).

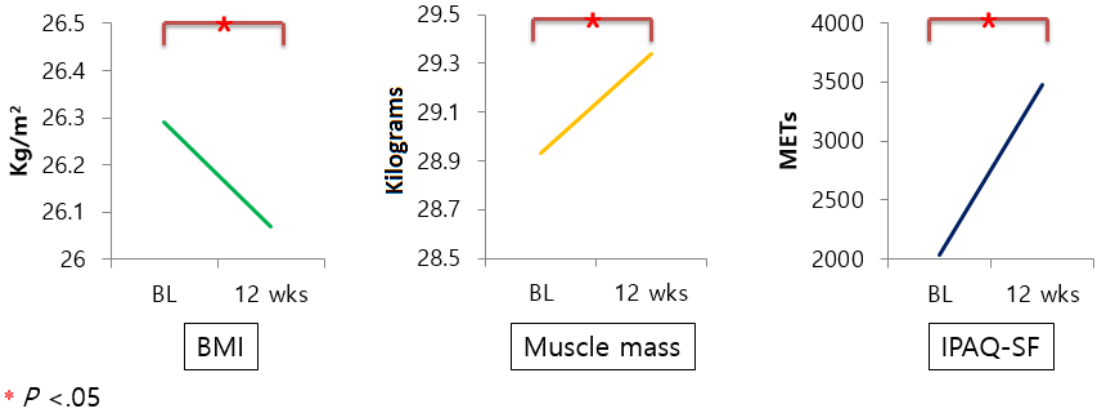
Table 3. Changes in body composition and self-reported physical activity.

Measure	Baseline, mean (SD)	12 weeks, mean (SD)	*P* value
Body mass index (kg/m^2)	25.39 (3.00)	25.57 (3.08)	.08
Body fat (%)	26.29 (8.01)	26.07 (7.86)	.51
Muscle mass (kg)	28.98 (5.15)	29.34 (5.31)	.03
IPAQ-SF[a] (METs[b])	2031.95 (2236.60)	3479.71 (2640.08)	.01

[a]IPAQ-SF: International Physical Activity Questionnaire-Short Form.

[b]METs: metabolic equivalents.

Figure 4. Changes in body composition and self-reported physical activity. BL: baseline; BMI: body mass index; IPAQ-SF: International Physical Activity Questionnaire-Short Form; METs: metabolic equivalents.



Biochemical Profile and Complications

Serum levels of aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase, and gamma glutamyl transferase, which indicate liver function, showed no significant change after 12 weeks of exercise. Insignificant results were seen in albumin and creatinine, which are synthesized in liver cells, and in total bilirubin and prothrombin time (ie, international normalized ratio). No significant difference was observed in osteocalcin, which is inversely related to nonalcoholic fatty liver disease (NAFLD); cholesterol, which increases in NAFLD; or complete blood count. During 12 weeks of exercise, no complications were reported through the in-app chat service, including hepatic decompensation in the form of variceal hemorrhage, ascites, jaundice, or hepatic encephalopathy.

Quality of Life

Table 4 presents the changes in QoL after 12 weeks of monitored, individualized exercise via our mHealth app. Based on the EORTC-QLQ-C30, the most common symptom was fatigue, followed by dyspnea and insomnia. After 12 weeks of individually prescribed exercise through the mHealth app, all symptoms improved insignificantly, except pain, which decreased significantly (P=.04). Global health status and all functional scales trended toward improvement after 12 weeks of the intervention, though without statistical significance.

Table 4. Changes in quality of life.

EORTC-QLQ-C30[a] item	Score at baseline, mean (SD)	Score at 12 weeks, mean (SD)	*P* value
Global health status and quality of life[b]	72.50 (17.52)	74.44 (17.63)	.43
Functional scale [b]			
Physical functioning	85.77 (10.31)	87.55 (11.30)	.43
Role functioning	87.22 (13.62)	89.44 (16.65)	.38
Emotional functioning	84.16 (18.48)	87.50 (12.90)	.28
Cognitive functioning	80.55 (13.19)	86.11 (11.64)	.06
Social functioning	85.55 (24.65)	93.33 (12.06)	.08
Symptom scale or single item[c]			
Fatigue	25.18 (12.69)	23.33 (15.25)	.48
Nausea and vomiting	2.22 (5.76)	2.17 (6.31)	.66
Pain	10.55 (14.17)	6.11 (11.14)	.04
Dyspnea	22.22 (26.74)	17.77 (24.34)	.35
Insomnia	15.55 (22.71)	14.44 (20.86)	.80
Appetite loss	6.66 (16.14)	5.55 (12.63)	.71
Constipation	12.00 (17.83)	11.11 (15.98)	.80
Diarrhea	12.00 (17.83)	11.11 (15.98)	.71
Financial difficulties	13.33 (25.67)	7.77 (16.80)	.20

[a]EORTC-QLQ-C30: European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire C30.

[b]Higher scores imply positive results.

[c]Lower scores imply positive results.

Discussion

Principal Findings

In this study, we found that it was safe and effective for compensated HCC patients who have completed anticancer therapy to undergo 12 weeks of individually prescribed rehabilitation exercises using our mHealth app and interconnected IoT wearable device. Our surveillance system, composed of the wearable IoT device and real-time communication chat service with health care professionals, found no complications or biochemical deterioration during the 12 weeks of the intervention. Compared to baseline, statistically significant improvements were found in the physical fitness measures (ie, grip strength, 30-second chair stand test, and 6MWT), body composition (ie, muscle mass), self-reported amount of physical activity (ie, IPAQ-SF), and pain. All symptoms trended toward improvement in the QoL scales (ie, EORTC-QLQ C30) after 12 weeks of the intervention. Compliance with the aerobic and strengthening exercises decreased slightly but was maintained overall throughout the intervention period. To the best of our knowledge, this was the first study to use an mHealth app and interconnected IoT device for individually prescribed rehabilitation exercises in compensated HCC patients after anticancer treatment.

Growing interest had led to many studies investigating the association of cancer and exercise. Increased risks for several solid cancers, including breast, endometrial, prostate, and colorectal, were inversely associated with physical activity [21].

Recent studies indicate that even after diagnosis with those solid cancers, regular exercise resulted in more survival years and a lower tumor recurrence rate [22-24]. The biochemical mechanisms of such findings are still being discovered, but exercise-dependent regulation of the tumor microenvironment is currently held to be the primary mechanism. In experimental models, chronic exposure to exercise stimulated interorgan signaling through the complex control of hormones, cytokines, and growth factors. Consequently, reprogramming of the systemic milieu was stimulated, resulting in regulation of the tumor microenvironment through angiogenesis, immune regulation, and metabolism [25].

The potential benefit of exercise on liver disease was first studied in NAFLD, of which the key pathophysiological mechanism is insulin resistance [26]. In terms of HCC, a Taiwanese cohort study [27], which was later confirmed by the National Institutes of Health [28], found that a degree of physical activity correlated with a decline in HCC risk. The antitumoral effect of exercise on HCC is mainly associated with decreases in body weight, insulin resistance, and chronic inflammation [29]. In a previous experimental study with a rodent model of nonalcoholic steatohepatitis, regular exercise inhibited HCC development by stimulating adenosine monophosphate-activated protein kinase and inhibiting mammalian target of rapamycin complex 1 [7].

However, only a few studies have considered the potential therapeutic effects of exercise after the development of HCC. One Japanese study had HCC patients work out at their anaerobic thresholds from 1 month preoperation until 6 months postoperation. The exercise group showed decreased whole-body mass, fat mass, and fasting serum insulin; they even showed improvement in the anaerobic threshold, peak oxygen consumption, and insulin resistance [30]. Contrary to our intervention period of 12 weeks, the Japanese study's program began before hepatectomy and lasted for 6 months postoperation. Also, that study's exercise intervention was targeted to the anaerobic threshold and consisted only of stretching and walking, whereas our study's exercise protocol combined stretching, aerobic, and muscle-strengthening exercises. Both studies adjusted the exercise program during the intervention. Most importantly, compared to the Japanese patients who visited the study center three times a week for 60-minute exercise sessions with an exercise trainer, our participants exercised freely outside the hospital via the mHealth program. In our study, exercise compliance and safety were monitored via the mHealth devices.

Another study reported that a median of 13 days of in-hospital exercise among HCC patients with underlying CLD did not worsen the Child-Pugh class, maintained the 6MWT distance, and significantly improved heart rate variability [6]. Compared to the thorough exercise program for upper and lower extremities in our study protocol, the exercise program of that previous study contained stretching, strengthening, and balance training for only the lower extremities. Unlike our study, the exercises in the previous study were not tailored according to individual fitness levels nor adjusted during the intervention period. Also, those exercises were initiated 1 day after anticancer treatment for HCC. Though the previous study evaluated 6MWT and biochemical profiles by blood sampling, the relatively short exercise period (ie, median 7.5 days) of that in-hospital study was not long enough to show physiological or biological changes.

Contrary to malignancies that occur without any underlying condition, HCC commonly arises in cirrhotic liver or viral hepatitis. Most HCC patients suffer from CLD for several years before developing HCC, so their QoL is low and worsens after cancer treatment. One global survey study found that among HCC patients, 81% receiving sorafenib (ie, tyrosine kinase inhibitor), 45% receiving selective internal radiation therapy, and 32% receiving transarterial chemoembolization reported impaired QoL [31]. Our unpublished previous survey study also found that CLD patients, including those with HCC, reported a low level of health-related QoL and physical activity, being minimally active. Boosting the low physical fitness levels of chronically ill HCC patients is a complex process, involving both motivating and safely leading these patients to exercise properly.

As the chronicity of the underlying CLD in HCC patients deteriorates their general physical activity levels, we used an mHealth app via a mobile phone and interconnected wearable IoT device. This system allowed close monitoring of patients' vital signs and exercise compliance, which allowed all participants to complete the 12-week exercise program without any complications or biochemical deterioration. Despite increasing interest and studies on mHealth for cancer patients, studies involving the use of IoT devices among cancer patients are scarce. A few studies using wearable IoT devices and mHealth apps were conducted to promote physical activity in breast cancer survivors and childhood cancer survivors [32-34]. Only two studies used an IoT device among cancer patients under treatment: one is our previous study involving colorectal cancer patients receiving chemotherapy, and the other study estimated the symptom severity of chemotherapy with a wearable IoT device in gastrointestinal cancer patients [4,35]. We are the first to apply mHealth with an IoT device to HCC patients. According to the satisfaction survey after 12 weeks of our mHealth program for HCC patients, 84% of participants reported medium-to-high satisfaction with the mHealth program. A total of 87% of participants wanted to continue to use the program after the study ended.

This study has several limitations. First, the study population was small and the intervention period was short. In most of the previously conducted feasibility or pilot studies regarding new protocols of exercise interventions among a cancer population, 10-30 patients were enrolled [36-39]. Since this was the first study to apply an IoT-based individualized exercise program among compensated HCC patients who had completed therapy, we considered that the final participant number of 31 was sufficient. The intervention period was selected based on previous interventional studies using mobile apps among cancer patients. Most studies were conducted for 12 weeks or 3 months [11,18,40]. Further studies with large sample sizes and longitudinal intervention periods will support the results of this study. Second, the completion rate of this study was low. The completion rate in exercise intervention studies of cancer populations differs based on the traits of the population, such

as dominant gender, mean age, severity of cancer, general condition, needs of exercise, and type of ongoing anticancer therapy. For example, one study that compared the physical fitness effect of mHealth and conventional exercise using a brochure among breast cancer patients—all female patients, mean age 50.3 years, and 12-week follow-up—showed a 95.2% completion rate [18]. In contrast, 73.9% of participants diagnosed with prostate cancer—all male patients, mean age 68.4 years, and 12-week follow-up—completed the 12-week course of a home-based *exergaming* intervention [41]. The completion rate of 84% (31/37) in this study—84% male patients, mean age 56.7 years, and 12-week follow-up—is not lower than that of the previous prostate cancer exercise intervention. Also, without a previous study of a 12-week exercise intervention among HCC patients, it is difficult to confirm that the completion rate of our study is low. Third, we used no control group in this study. Inclusion of a sex- and age-matched control group could have strengthened our results. However, as this was the first study to apply an mHealth-based exercise program to HCC patients, comparing the outcome measures of the initial, 6-week, and 12-week visits of the mHealth exercise group without a control group yielded significant results. Finally, because this study was conducted among compensated HCC patients after cancer therapy, it would be inappropriate to generalize these results. Further studies with a larger sample size and a randomized controlled study involving HCC patients with diverse medical statuses are needed.

Conclusions

In this study, compensated HCC patients after therapy underwent 12 weeks of comprehensive cancer care through an mHealth app. The app provided general medical information on HCC, nutritional information, real-time communication with a health care provider, and an individually prescribed rehabilitation exercise program. The exercise program was monitored through an IoT device. The intervention significantly improved physical fitness, body composition, and self-reported physical activity without any complications or biochemical deterioration. We found it safe and effective for compensated HCC patients after cancer treatment to exercise using the mobile app and interconnected IoT wearable device. This advanced technology allows effective and practical, patient-centered, supervised home exercise therapy to be an alternative to conventional, unfeasible, hospital-oriented exercise programs.

Conflicts of Interest

None declared.

References

1. Wirtz P, Baumann FT. Physical activity, exercise and breast cancer: What is the evidence for rehabilitation, aftercare, and survival? A review. Breast Care (Basel) 2018 Apr;13(2):93-101 [FREE Full text] [doi: 10.1159/000488717] [Medline: 29887785]
2. Okumatsu K, Tsujimoto T, Wakaba K, Seki A, Kotake R, Yamauchi T, et al. Effects of a combined exercise plus diet program on cardiorespiratory fitness of breast cancer patients. Breast Cancer 2019 Jan;26(1):65-71. [doi: 10.1007/s12282-018-0889-x] [Medline: 29992486]
3. Campos C, Sotomayor P, Jerez D, González J, Schmidt CB, Schmidt K, et al. Exercise and prostate cancer: From basic science to clinical applications. Prostate 2018 Jun;78(9):639-645. [doi: 10.1002/pros.23502] [Medline: 29569731]
4. Cheong IY, An SY, Cha WC, Rha MY, Kim ST, Chang DK, et al. Efficacy of mobile health care application and wearable device in improvement of physical performance in colorectal cancer patients undergoing chemotherapy. Clin Colorectal Cancer 2018 Jun;17(2):e353-e362. [doi: 10.1016/j.clcc.2018.02.002] [Medline: 29551558]
5. Mittal S, El-Serag HB. Epidemiology of hepatocellular carcinoma: Consider the population. J Clin Gastroenterol 2013 Jul;47 Suppl:S2-S6 [FREE Full text] [doi: 10.1097/MCG.0b013e3182872f29] [Medline: 23632345]
6. Koya S, Kawaguchi T, Hashida R, Goto E, Matsuse H, Saito H, et al. Effects of in-hospital exercise on liver function, physical ability, and muscle mass during treatment of hepatoma in patients with chronic liver disease. Hepatol Res 2017 Mar;47(3):E22-E34. [doi: 10.1111/hepr.12718] [Medline: 27062043]
7. Piguet A, Saran U, Simillion C, Keller I, Terracciano L, Reeves HL, et al. Regular exercise decreases liver tumors development in hepatocyte-specific PTEN-deficient mice independently of steatosis. J Hepatol 2015 Jun;62(6):1296-1303. [doi: 10.1016/j.jhep.2015.01.017] [Medline: 25623824]
8. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: A systematic review. J Med Internet Res 2016 May 16;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
9. Sankaran S, Dendale P, Coninx K. Evaluating the impact of the HeartHab app on motivation, physical activity, quality of life, and risk factors of coronary artery disease patients: Multidisciplinary crossover study. JMIR Mhealth Uhealth 2019 Apr 04;7(4):e10874 [FREE Full text] [doi: 10.2196/10874] [Medline: 30946021]
10. Wu X, Guo X, Zhang Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: Systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
11. Geng Z, Wu F, Zhang Y, Wang J, Wang J, Liu H, et al. Mobile physical activity intervention for breast cancer patients during chemotherapy. Stud Health Technol Inform 2018;250:236. [Medline: 29857447]

12. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: A mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
13. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: A systematic content analysis. J Med Internet Res 2014 Aug 15;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
14. Enright PL, Sherrill DL. Reference equations for the six-minute walk in healthy adults. Am J Respir Crit Care Med 1998 Nov;158(5 Pt 1):1384-1387. [doi: 10.1164/ajrccm.158.5.9710086] [Medline: 9817683]
15. 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]
16. Sternby Eilard M, Hagström H, Mortensen KE, Wilsgaard T, Vagnildhaug OM, Dajani O, et al. Quality of life as a prognostic factor for survival in hepatocellular carcinoma. Liver Int 2018 May;38(5):885-894. [doi: 10.1111/liv.13593] [Medline: 28941130]
17. Fayers PA, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. EORTC QLQ-C30 Scoring Manual. 3rd edition. Brussels, Belgium: European Organisation for Research and Treatment of Cancer (EORTC); 2001. URL: https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf [accessed 2020-01-29]
18. Uhm KE, Yoo JS, Chung SH, Lee JD, Lee I, Kim JI, et al. Effects of exercise intervention in breast cancer patients: Is mobile health (mHealth) with pedometer more effective than conventional program using brochure? Breast Cancer Res Treat 2017 Feb;161(3):443-452. [doi: 10.1007/s10549-016-4065-8] [Medline: 27933450]
19. Millor N, Lecumberri P, Gómez M, Martínez-Ramírez A, Izquierdo M. An evaluation of the 30-s chair stand test in older adults: Frailty detection based on kinematic parameters from a single inertial unit. J Neuroeng Rehabil 2013 Aug 01;10:86 [FREE Full text] [doi: 10.1186/1743-0003-10-86] [Medline: 24059755]
20. Korea Sports Promotion Foundation. National fitness 100 URL: http://nfa.kspo.or.kr/front/control/con0901_list.do [accessed 2020-01-28]
21. Kerr J, Anderson C, Lippman SM. Physical activity, sedentary behaviour, diet, and cancer: An update and emerging new evidence. Lancet Oncol 2017 Aug;18(8):e457-e471. [doi: 10.1016/S1470-2045(17)30411-4] [Medline: 28759385]
22. Chen X, Lu W, Zheng W, Gu K, Matthews CE, Chen Z, et al. Exercise after diagnosis of breast cancer in association with survival. Cancer Prev Res (Phila) 2011 Sep;4(9):1409-1418 [FREE Full text] [doi: 10.1158/1940-6207.CAPR-10-0355] [Medline: 21795422]
23. Friedenreich CM, Wang Q, Neilson HK, Kopciuk KA, McGregor SE, Courneya KS. Physical activity and survival after prostate cancer. Eur Urol 2016 Oct;70(4):576-585 [FREE Full text] [doi: 10.1016/j.eururo.2015.12.032] [Medline: 26774959]
24. Song M, Chan AT. The potential role of exercise and nutrition in harnessing the immune system to improve colorectal cancer survival. Gastroenterology 2018 Sep;155(3):596-600 [FREE Full text] [doi: 10.1053/j.gastro.2018.07.038] [Medline: 30076837]
25. Koelwyn GJ, Quail DF, Zhang X, White RM, Jones LW. Exercise-dependent regulation of the tumour microenvironment. Nat Rev Cancer 2017 Sep 25;17(10):620-632. [doi: 10.1038/nrc.2017.78] [Medline: 28943640]
26. Kirwan JP, Solomon TP, Wojta DM, Staten MA, Holloszy JO. Effects of 7 days of exercise training on insulin sensitivity and responsiveness in type 2 diabetes mellitus. Am J Physiol Endocrinol Metab 2009 Jul;297(1):E151-E156 [FREE Full text] [doi: 10.1152/ajpendo.00210.2009] [Medline: 19383872]
27. Wen C, Lin J, Yang Y, Tsai M, Tsao C, Etzel C, et al. Hepatocellular carcinoma risk prediction model for the general population: The predictive power of transaminases. J Natl Cancer Inst 2012 Oct 17;104(20):1599-1611 [FREE Full text] [doi: 10.1093/jnci/djs372] [Medline: 23073549]
28. Behrens G, Matthews CE, Moore SC, Freedman ND, McGlynn KA, Everhart JE, et al. The association between frequency of vigorous physical activity and hepatobiliary cancers in the NIH-AARP Diet and Health Study. Eur J Epidemiol 2013 Jan;28(1):55-66 [FREE Full text] [doi: 10.1007/s10654-013-9767-1] [Medline: 23354983]
29. Saran U, Humar B, Kolly P, Dufour J. Hepatocellular carcinoma and lifestyles. J Hepatol 2016 Jan;64(1):203-214. [doi: 10.1016/j.jhep.2015.08.028] [Medline: 26341826]
30. Kaibori M, Ishizaki M, Matsui K, Nakatake R, Yoshiuchi S, Kimura Y, et al. Perioperative exercise for chronic liver injury patients with hepatocellular carcinoma undergoing hepatectomy. Am J Surg 2013 Aug;206(2):202-209. [doi: 10.1016/j.amjsurg.2012.07.035] [Medline: 23374372]
31. Gill J, Baiceanu A, Clark PJ, Langford A, Latiff J, Yang P, et al. Insights into the hepatocellular carcinoma patient journey: Results of the first global quality of life survey. Future Oncol 2018 Jul;14(17):1701-1710. [doi: 10.2217/fon-2017-0715] [Medline: 29543521]
32. Ritvo P, Obadia M, Santa Mina D, Alibhai S, Sabiston C, Oh P, et al. Smartphone-enabled health coaching intervention (iMOVE) to promote long-term maintenance of physical activity in breast cancer survivors: Protocol for a feasibility pilot randomized controlled trial. JMIR Res Protoc 2017 Aug 24;6(8):e165 [FREE Full text] [doi: 10.2196/resprot.6615] [Medline: 28838886]

33. Mendoza JA, Baker KS, Moreno MA, Whitlock K, Abbey-Lambertz M, Waite A, et al. A Fitbit and Facebook mHealth intervention for promoting physical activity among adolescent and young adult childhood cancer survivors: A pilot study. Pediatr Blood Cancer 2017 Dec;64(12):1. [doi: 10.1002/pbc.26660] [Medline: 28618158]
34. Valle CG, Deal AM, Tate DF. Preventing weight gain in African American breast cancer survivors using smart scales and activity trackers: A randomized controlled pilot study. J Cancer Surviv 2017 Feb;11(1):133-148 [FREE Full text] [doi: 10.1007/s11764-016-0571-2] [Medline: 27631874]
35. Low CA, Dey AK, Ferreira D, Kamarck T, Sun W, Bae S, et al. Estimation of symptom severity during chemotherapy from passively sensed data: Exploratory study. J Med Internet Res 2017 Dec 19;19(12):e420 [FREE Full text] [doi: 10.2196/jmir.9046] [Medline: 29258977]
36. Brown M, Murphy M, McDermott L, McAneney H, O'Sullivan JM, Jain S, et al. Exercise for advanced prostate cancer: A multicomponent, feasibility, trial protocol for men with metastatic castrate-resistant prostate cancer (EXACT). Pilot Feasibility Stud 2019;5:102 [FREE Full text] [doi: 10.1186/s40814-019-0486-6] [Medline: 31428443]
37. Egegaard T, Rohold J, Lillelund C, Persson G, Quist M. Pre-radiotherapy daily exercise training in non-small cell lung cancer: A feasibility study. Rep Pract Oncol Radiother 2019;24(4):375-382. [doi: 10.1016/j.rpor.2019.06.003] [Medline: 31289452]
38. Yee J, Davis GM, Hackett D, Beith JM, Wilcken N, Currow D, et al. Physical activity for symptom management in women with metastatic breast cancer: A randomized feasibility trial on physical activity and breast metastases. J Pain Symptom Manage 2019 Dec;58(6):929-939. [doi: 10.1016/j.jpainsymman.2019.07.022] [Medline: 31374368]
39. Zhang X, McClean D, Ko E, Morgan M, Schmitz K. Exercise among women with ovarian cancer: A feasibility and pre-/post-test exploratory pilot study. Oncol Nurs Forum 2017 May 01;44(3):366-374. [doi: 10.1188/17.ONF.366-374] [Medline: 28635971]
40. Xiaosheng D, Xiangren Y, Shuyuan H, Dezong G, Mengyao C, Meng D. The effects of combined exercise intervention based on Internet and social media software for postoperative patients with breast cancer: Study protocol for a randomized controlled trial. Trials 2018 Sep 06;19(1):477 [FREE Full text] [doi: 10.1186/s13063-018-2857-3] [Medline: 30189899]
41. Villumsen BR, Jorgensen MG, Frystyk J, Hørdam B, Borre M. Home-based 'exergaming' was safe and significantly improved 6-min walking distance in patients with prostate cancer: A single-blinded randomised controlled trial. BJU Int 2019 Oct;124(4):600. [doi: 10.1111/bju.14782] [Medline: 31012238]

Abbreviations

6MWT: 6-minute walk test
ANOVA: analysis of variance
BMI: body mass index
CLD: chronic liver disease
EORTC-QLQ-C30: European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire C30
HCC: hepatocellular carcinoma
IoT: Internet of Things
IPAQ-SF: International Physical Activity Questionnaire-Short Form
METs: metabolic equivalents
mHealth: mobile health
mUICC: modified Union for International Cancer Control
NAFLD: nonalcoholic fatty liver disease
QoL: quality of life

Edited by G Eysenbach; submitted 18.04.19; peer-reviewed by L Scudeller, Y Tian; comments to author 03.10.19; revised version received 01.11.19; accepted 19.12.19; published 11.03.20.

Please cite as:
Kim Y, Seo J, An SY, Sinn DH, Hwang JH
Efficacy and Safety of an mHealth App and Wearable Device in Physical Performance for Patients With Hepatocellular Carcinoma: Development and Usability Study
JMIR Mhealth Uhealth 2020;8(3):e14435
URL: http://mhealth.jmir.org/2020/3/e14435/
doi:10.2196/14435
PMID:32159517

Review

Standards for Mobile Health–Related Apps: Systematic Review and Development of a Guide

Pere Llorens-Vernet[1], MSc; Jordi Miró[1], PhD

Unit for the Study and Treatment of Pain-ALGOS, Department of Psychology, Universitat Rovira i Virgili, Tarragona, Spain

Corresponding Author:
Jordi Miró, PhD
Unit for the Study and Treatment of Pain-ALGOS
Department of Psychology
Universitat Rovira i Virgili
Carretera de Valls, s/n
Tarragona, 43007
Spain
Phone: 34 977558179
Email: jordi.miro@urv.cat

Abstract

Background: In recent years, the considerable increase in the number of mobile health (mHealth) apps has made health care more accessible and affordable for all. However, the exponential growth in mHealth solutions has occurred with almost no control or regulation of any kind. Despite some recent initiatives, there is still no specific regulation procedure, accreditation system, or standards to help the development of the apps, mitigate risks, or guarantee quality.

Objective: The main aim of this study was to propose a set of criteria for mHealth-related apps on the basis of what is available from published studies, guidelines, and standards in the various areas that are related to health app development.

Methods: We used three sources of information to identify the most important criteria. First, we conducted a systematic review of all the studies published on pain-related apps. Second, we searched for health app recommendations on the websites of professional organizations. Third, we looked for standards governing the development of software for medical devices on the specialized websites of regulatory organizations. Then, we grouped and subsumed the criteria we had identified on the basis of their shared characteristics. Finally, the comprehensibility and perceived importance of the resulting criteria were evaluated for face validity with a group of 18 stakeholders.

Results: We identified a total of 503 criteria from all sources, which, after close analysis, were grouped into eight different categories, including 36 important criteria for health apps. The resulting categories were *usability, privacy, security, appropriateness and suitability, transparency and content, safety, technical support and updates*, and *technology*. The results of the preliminary analysis showed that the criteria were mostly understood by the group of stakeholders. In addition, they perceived all of them as important.

Conclusions: This set of criteria can help health care providers, developers, patients, and other stakeholders to guide the development of mHealth-related apps and, potentially, to measure the quality of an mHealth app.

(JMIR Mhealth Uhealth 2020;8(3):e13057) doi:10.2196/13057

KEYWORDS

mHealth; mobile apps; review; medical device; standards

Introduction

Background

Public health care systems worldwide are facing major challenges (eg, a shortage of resources and a steady increase in demand), which can make them increasingly unsustainable [1]. It is in this environment that what is known as mobile health (mHealth) is proving to be of key importance [2-4]. In the last few years, mHealth has undergone considerable development because of its potential to make health care more accessible and affordable for all [2,5-7].

However, mHealth solutions have grown exponentially with almost no control or regulation of any kind. In fact, very few of the health apps available have undergone a thorough

validation process, and this causes a lack of confidence among health professionals [8,9]. For example, a recent review of the mobile apps available for chronic pain—which is one of the most prevalent health problems, with an enormous economic cost to individuals, families, and society [10]—highlighted that of the 283 apps available at the time, just a handful had undergone usability and validity tests [7]. This situation has been identified as preventing the field from improving and advancing [9].

In this so-called *strategic field*, progress depends not only on what each research group is doing but also on developing general standards and improving certification procedures [11,12]. There are some local and international initiatives to help in this process. For example, Catalonia approved a strategic action plan to support the development of mHealth (ie, *The Mobility Master Plan: mHealth solutions* [13]), which includes *AppSalut* [14], an accreditation system and guide [15] created to certify the quality of health– and social-related apps. At the international level, the European Commission published a Green Paper on mHealth [16] and launched a public consultation to identify potential barriers to and problems in the development of mHealth. Despite these initiatives, there is still no specific regulation procedure, accreditation system, or standards to help the development of apps, mitigate risks, and guarantee quality.

Therefore, the certification process is weighed down by the lack of clear standards to guide users through the different stages of the process. This is a problem not only for the safety of end users (ie, patients and health care professionals) but also for professional developers. Clearly, having a set of common criteria would be instrumental in helping the field to make progress in a consensual way and overcome potential risks for all stakeholders. There has been one recent attempt to develop a rating scale for mobile apps that could be used to help overcome this problem. Stoyanov et al [17] developed a scale (Mobile App Rating Scale [MARS]) to classify and rate the quality of mHealth apps. This scale was on the basis of a review of the papers published between 2000 and 2013, which contained explicit app-related quality rating criteria. However, this scale was created from a very narrow perspective for assessing already developed apps. That is to say, although the authors used information from studies on existing mobile apps, they failed to include information from other relevant sources that had been used, which would have increased the reliability and validity of their work (eg, standards governing the development of software for health or medical devices). Therefore, this scale does not seem to be suited for use by all stakeholders. Some more recent attempts to provide alternatives to assess mHealth apps also share some of these weaknesses (eg, developed for one specific group of stakeholders and using one specific source of information) [18,19].

Objectives

The general aim of this study, then, is to go beyond what is already available and provide a standard for mHealth-related apps by studying the published studies, guidelines, and standards available in the field of health app development. In particular, we want to identify a set of criteria that are used, and which of these are strategic, so that they can be recommended and integrated into a general standard (ie, a guide) that can help the field move forward on solid grounds.

Methods

Procedure

We used three strategies to identify criteria. First, we conducted a systematic review of all the studies published on pain-related apps. Second, we searched the websites of professional organizations. Finally, we analyzed the standards governing the development of software for medical devices. Although these regulations are not specific for health apps, they can provide information of interest and complement the information collected.

Information From the Systematic Review

For the systematic review, to address an otherwise unmanageable amount of information, we limited our search to mobile apps related to pain, one of the most prevalent health problems causing millions of visits to health care professionals. In so doing, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis guidelines [20] and searched the following databases: Web of Science, Scopus, PubMed, and ScienceDirect. We used the search terms (pain OR *ache) AND (smartphone* OR mobile) AND (app OR apps), and also hand searched the reference lists from relevant articles. Only peer-reviewed articles published in English or Spanish between 2008 (the release date of the first apps stores [21]) and December 2017 were included.

Information on Websites From Professional Organizations

The second strategy consisted of searching websites from professional organizations that had guidelines and recommendations for health apps. We decided to focus our search on those regions where the mHealth market is most significant and, therefore, where these regulations are most likely to be found. According to a forecast of revenues of the world mHealth market [22], in 2017, the main mHealth markets by regions were Europe, North America, and Asia–Pacific, representing 30%, 28%, and 30% of the world market, respectively. In addition, the 15 most attractive countries for digital health solutions in 2017 were located in the regions mentioned above [23]. Therefore, we limited our search to those three areas. Again, and to make the search feasible, we limited our analysis to those countries that are the main markets in each region. In Europe, we included the United Kingdom and Spain as they had the same importance in terms of mHealth markets; in North America, the United States; and in Asia–Pacific, Australia.

Information From Standards Governing the Development of Software for Medical Devices

Finally, in our search for information, we also searched for standards governing the development of software for medical devices on the specialized websites of regulatory organizations. To conduct this search, we also focused on the regions and countries where the mHealth market has been shown to be most significant, as described above. In this analysis, we focused the

search on those standards with criteria related to health apps, added to our list only those criteria that are specific to health apps, and left all others out of our scrutiny (eg, protection against radiation and chemical properties).

Development of a Common Set of Criteria and Categories

We first compiled a list of the criteria identified in (1) published studies, (2) guidelines, and (3) standards governing the development of software for medical devices. Next, we grouped the criteria in categories on the basis of their shared characteristics. That is, each criterion was closely analyzed to identify what its general purpose was (eg, the criterion *the functionality is adapted to the purpose of the app* was considered related to usability, and this opened a group or category that was labeled *usability*). All criteria underwent the same scrutiny. In the case that no category existed, a new one was created and labeled. If the category already existed, then the criterion was subsumed under that existing category. As a result of this analysis, we obtained a list of unique criteria classified into categories according to their similarity.

Preliminary Analysis of the Set of Criteria

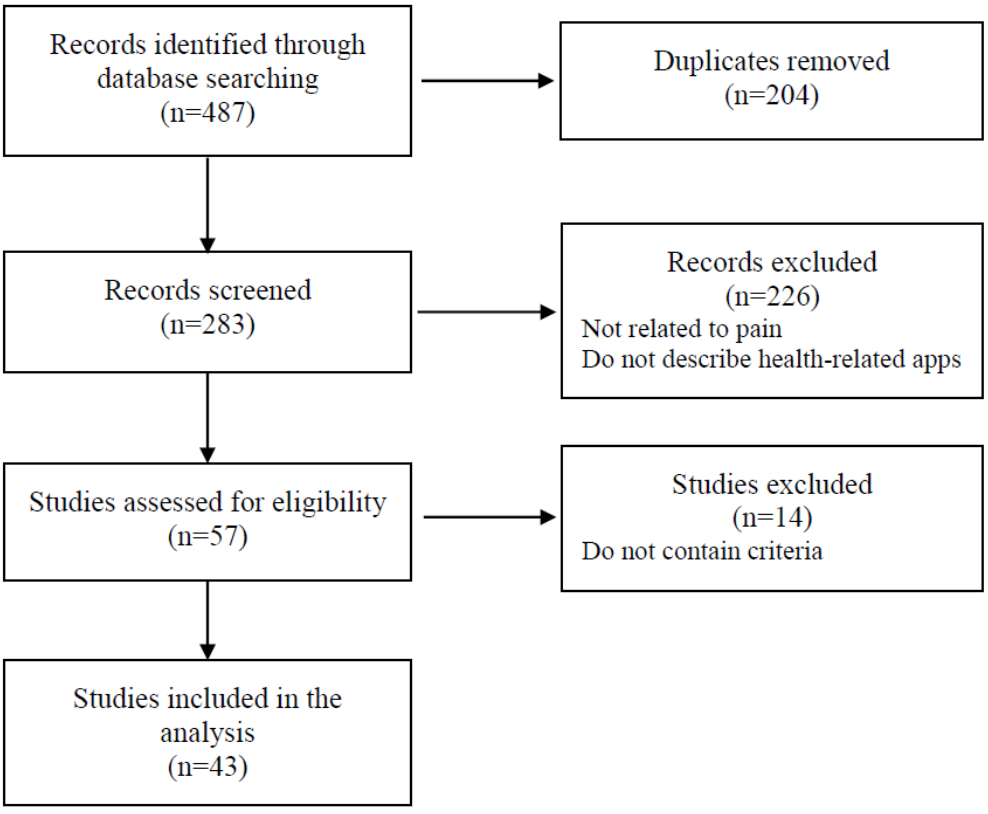
The resulting set of criteria underwent a preliminary analysis of their face validity by asking stakeholders to report on the comprehensibility and perceived importance of all the criteria. Specifically, in this analysis, we requested the collaboration of a group of individuals from different groups of stakeholders (ie, final users, potential patients, health care professionals, and developers or engineers). Final users or potential patients and health care professionals were approached by the authors while at the university hospital (while they were visiting for a health checkup and while at work, respectively). Engineers were professors or technicians working at the university. Before the participation of stakeholders, we first requested and obtained permission from the Ethics Committee of the School of Education Sciences and Psychology for the study procedures. Participants had to sign a consent form. All were asked to respond to two questions in relation to each criterion: (1) “Do you understand the criterion?” and (2) “How important is this criterion for a health-related mobile application?” The first one was responded with *yes*, *no*, or *partially* to the question, whereas the second one was to be responded by providing a number that best represented the importance of the criterion, between 0 (*not important at all*) to 10 (*utmost important*).

Results

Information From the Systematic Review

Our review of the scientific databases identified 283 nonduplicated papers. Of these, only 43 were of interest for our purposes. Studies that were not related to pain or that did not describe health-related apps were deemed irrelevant, and not included in the analysis (see Figure 1). In this search, 168 criteria were identified (the full list is provided in Multimedia Appendix 1).

Figure 1. Flowchart of systematic review selection process.



Information From Websites of Professional Organizations

Following the planned strategy, we found just 4 organizations that had developed guidelines and recommendations for health-related apps. Of these, 3 were of national coverage—Andalusian Agency for Healthcare Quality (Spain), TIC Salut Social Foundation (Spain), and National Health Service (United Kingdom)—and 1 of supranational or international coverage, the European Commission (European Union [EU]). No similar information was found in the other searched regions. From each of the guidelines, we collected only the criteria that were specifically related to mobile phone apps and discarded the criteria related to other technologies (eg, wearables and websites): Andalusian Agency for Healthcare Quality, 31 criteria; European Commission, 58 criteria; National Health Service, 78 criteria; and TIC Salut Social Foundation, 115 criteria (see Multimedia Appendix 1).

Information From Standards Governing the Development of Software For Medical Devices

As planned, in our search of the main mHealth markets, we also looked at the specialized websites of regulatory organizations and searched through standards in the regulations of medical devices. In so doing, we found just two standards that were of interest: (1) *Mobile Medical Applications: Guidance for Food and Drug* (the United States) and (2) *Regulation of medical software andmobile medical "apps"* (Australia). In this analysis, we added to our list only those criteria that were specific to health apps and left all others out of our scrutiny (eg, protection against radiation and chemical properties). We identified 42 and 11 criteria, respectively (see Multimedia Appendix 1).

Development of a Common Set of Criteria and Categories

Then, the set of criteria were grouped in categories according to their similarity. That is to say, the criteria of the same class were grouped and subsumed together (see Table 1), resulting in eight categories. The categories were the following: *usability* (this includes criteria that are related to user experience), which contained eight criteria; *privacy* (ie, criteria related to data protection, compliance with the law, and treatment of users' data), which contained six criteria; *security* (ie, criteria related to cybersecurity, encryption mechanisms for the storage and transmission of data, and measures against vulnerabilities), which contained four criteria; *appropriateness and suitability* (ie, criteria related to the adaptation of the app for the benefit of the targeted user), which contained three criteria; *transparency and content* (ie, criteria related to the sharing of information in relation to the development of the app), which contained five criteria; *safety* (ie, criteria related to the identification and prevention of harm to end users), which contained two criteria; *technical support and updates* (ie, criteria related to helping the user to solve problems in using the app), which contained four criteria; and *technology* (ie, criteria related to the proper functioning of the app), which contained four criteria (see Table 1).

Preliminary Analysis of the Set of Criteria

A total of 18 individuals participated: 7 final users or potential patients, 6 health care professionals, and 5 developers, all of whom were approached and consented. Participants' age ranged from 18 to 53 years, with an equal distribution of females and males in the sample. At the time of participation, all were attending school or working. They all had experience with mobile phones and in using mobile apps.

The results of this analysis are summarized in Table 1, which includes information about the percentage of participants within each group that understood the criteria, and the mean of the perceived importance of each one.

The criteria were understood by most of the participants in the three groups. However, at least one participant in one or more groups reported being unsure about the exact meaning. All the issues were related to the use of technical vocabulary or lack of some very specific (technical) knowledge; nevertheless, with additional explanation, the issues were solved. In addition, all criteria were perceived as important; 7 (on a 0-10 numerical rating scale) was the lowest rating received by any criterion, and most ratings were between 8 and 10 (Table 1 summarizes the information).

Table 1. Comprehensibility and perceived importance of the criterion by stakeholders.

Category and criterion	Comprehension, n (%)			Perceived importance (0-10)		
	Patients (N=7)	Clinicians (N=6)	Engineers (N=5)	Patients	Clinicians	Engineers
Usability						
The app has been tested by potential users before being made available to the public.	7 (100)	5 (83)	5 (100)	8.9	8.6	9.4
It has instructions or some kind of assistance for use.	6 (86)	6 (100)	5 (100)	10	8.5	7
It is easy to use (ie, navigation is intuitive).	7 (100)	6 (100)	5 (100)	9.6	9.3	9
It follows the recommendations, patterns, and directives in the official manuals of the different operating systems (Android, iOS, or others).	7 (100)	5.5 (92)	4 (80)	7.5	7	7.9
The interface design follows the same pattern. That is, all graphic elements (typographies, icons, and buttons) have a consistent appearance. The function of each element (navigation menu, lists, and photo gallery) is clearly identified.	7 (100)	6 (100)	5 (100)	9	8.3	8.8
The functionality is adapted to the purpose of the app.	7 (100)	5.5 (92)	5 (100)	9.6	8.5	8.8
The information of the app must be able to be accessed in the shortest possible time. All users must be able to access all resources regardless of their capabilities.	7 (100)	5.5 (92)	5 (100)	8.3	8.6	8.4
The app can be consulted in more than one language. All languages adapt appropriately to the content interface.	7 (100)	5.5 (92)	5 (100)	8.3	7.4	7.2
Privacy						
The app gives information about the terms and conditions of purchases in the app and personal data recorded.	7 (100)	6 (100)	5 (100)	9.5	8.5	9
It gives information about the kind of user data to be collected and the reason (the app must only ask for user data that is essential for the app to operate). It gives information about access policies and data treatment and ensures the right of access to recorded information. It describes the maintenance policy and the data erasure procedure. It gives information about possible commercial agreements with third parties.	7 (100)	5.5 (92)	5 (100)	9.6	8.8	9.4
It guarantees the privacy of the information recorded. It requires users to give their express consent. It warns of the risks of using the app.	7 (100)	6 (100)	5 (100)	9.9	9.2	9
It tells users when it accesses other resources of the device, such as their accounts or their social network profile.	7 (100)	6 (100)	5 (100)	9.3	8.3	9.4
It takes measures to protect minors in accordance with the current legislation.	7 (100)	5.5 (92)	5 (100)	8.7	9.7	8.2
Confidential user data are protected and anonymized, and there is a privacy mechanism so that users can control their data.	7 (100)	6 (100)	5 (100)	9.6	9.2	9.4
Security						
The app has encryption mechanisms for storing, collecting, and exchanging information. It has password management mechanisms.	7 (100)	6 (100)	5 (100)	9.9	8.3	8.6
The cloud services used have the relevant security measures. It states the terms and conditions of cloud services.	7 (100)	6 (100)	5 (100)	9.5	8.2	8.4
The authorization and authentication mechanisms protect the users' credentials and gives access to their data. It limits access to data that is only necessary for the user.	7 (100)	6 (100)	5 (100)	9.6	7.5	9.6
It detects and identifies cybersecurity vulnerabilities, possible threats, and the risk of being exploited. It applies the appropriate security measures to cybersecurity vulnerabilities in the face of possible threats.	7 (100)	5.5 (92)	5 (100)	9.3	8.3	8.4
Appropriateness and suitability						
The end users for whom the app is designed are explicitly indicated or actually intuitable (the name identifies the app) to the audience to whom it is set out.	7 (100)	6 (100)	5 (100)	8.2	7.7	8.4

Category and criterion	Comprehension, n (%)			Perceived importance (0-10)		
	Patients (N=7)	Clinicians (N=6)	Engineers (N=5)	Patients	Clinicians	Engineers
The benefits and advantages of using the app are explained.	7 (100)	6 (100)	5 (100)	9	7	7.5
The app has been validated or created by experts (eg, a group of specialized professionals, a health organization, or a scientific society).	6.5 (93)	5.5 (92)	5 (100)	8	9.7	9
Transparency and content						
The app identifies the authors of the content and their professional qualifications.	7 (100)	6 (100)	5 (100)	8	8.5	8.4
It gives transparent information about the owners' identity and location.	7 (100)	6 (100)	5 (100)	7.2	8.2	8.4
It gives information about its sources of funding, promotion and sponsorship, and possible conflicts of interests. Any third parties or organizations who have contributed to the app development are clearly identified.	7 (100)	6 (100)	5 (100)	7.5	7	7
It uses scientific evidence to guarantee the quality of the content. It is based on ethical principles and values.	7 (100)	6 (100)	5 (100)	10	9.3	9
The sources of the information are indicated. Concise information is given about the procedure used to select the content.	7 (100)	6 (100)	5 (100)	8.2	7	8
Safety						
The possible risks to users are identified. Users are warned that the app does not intend to replace the services provided by a professional.	7 (100)	6 (100)	5 (100)	9.5	8.9	8.6
Potential risks for users caused by bad usage or possible adverse effects are explained.	7 (100)	6 (100)	5 (100)	8.5	8.5	8.6
Technical support and updates						
It gives a warning if updates modify or affect how the app functions. It gives a warning if updates can influence insensitive data.	7 (100)	6 (100)	5 (100)	8.5	7.2	7
Frequent security updates are guaranteed. Every time an update of a third-party component is published, the change is inspected, and the risk evaluated.	7 (100)	5.5 (92)	4.5 (90)	8.2	8.2	7
The frequency with which the content of the app is revised or updated is shown.	7 (100)	6 (100)	4.5 (90)	7.7	7	7
Users have support mechanisms (email, phone, and contact form) for solving doubts, problems, or issues related to the health content, and technical support.	7 (100)	6 (100)	5 (100)	9.2	9	8.4
Technology						
It works correctly. It does not fail during use (eg, blocks). Functions are correctly retrieved after context changes (eg, switch to another app and return), external interruptions (eg, incoming calls or messages), and switching off the terminal.	7 (100)	6 (100)	5 (100)	9.5	8.5	9
It does not waste resources excessively: battery, central processing unit, memory, data, or network.	7 (100)	6 (100)	5 (100)	8.9	7.8	8.2
It can work in flight mode and deal with network delays and any loss of connection.	7 (100)	5.5 (92)	5 (100)	7.9	7.2	7
It supports multiple versions of data structures or formats (eg, to support different operating systems).	6.5 (93)	5.5 (92)	4.5 (90)	8	7	7

Discussion

Principal Findings

To the best of our knowledge, this study is the first one to provide a guide to help with the design, development, and analysis of mHealth-related apps, in the form of a list of criteria and categories. This guide is based on an in-depth analysis of criteria that have been described in published studies on pain-related mHealth apps, guidelines, and best practices, as designated on the websites of professional and regulatory organizations from the most significant regions and countries of the world mHealth market.

In this study, we identified 36 criteria that are important to the design, development, and analysis of mHealth-related apps,

which were grouped and subsumed into eight categories according to their similarity: (1) *usability* (ie, the app must be adapted to the targeted population), (2) *privacy* (ie, compliance with the law and treatment of users' data), (3) *security* (ie, data protection, authorization mechanisms, and detection of vulnerability), (4) *appropriateness and suitability* (ie, the benefits and advantages for the end users are explained), (5) *transparency and content* (ie, scientific evidence and sources information), (6) *safety* (ie, the potentiality of risk to end users), (7) *technical support and updates* (ie, there is a policy about the maintenance of the app after it has been launched), and (8) *technology* (ie, the app works smoothly and does not fail abruptly).

In addition, this set of criteria underwent a test, and the preliminary data have shown that the criteria are understood by potential users. Furthermore, they have been reported to be of high importance by the group of stakeholders. Of particular importance (ie, a criterion that was valued as 9 or higher by all stakeholders groups on a 0-10 numerical rating scale) were the following: (1) *It is easy to use* (ie, navigation is intuitive); (2) *It guarantees the privacy of the information recorded. It requires users to give their express consent. It warns of the risks of using the app*; (3) *Confidential user data is protected and anonymized, and there is a privacy mechanism so that users can control their data*; and (4) *It uses scientific evidence to guarantee the quality of the contents. It is based on ethical principles and values*.

Our work improves previous proposals as it brings together information from a variety of internationally relevant sources (ie, research studies, data from websites of professional organizations, and standards governing the development of software for health or medical devices), whereas available ones have been developed narrowly, mostly using just one source (eg, studies on mobile apps [17]), sometimes using data of unknown scientific value (ie, mobile apps available on Web-based stores that have not undergone usability or validity studies [19]). This might be responsible, at least in part, for missing information in available guides. For example, in the case of the MARS [17], which is one of the most used rating systems, authors have failed to include some very basic items on their scale. Of particular concern are the issues of privacy and security of users' information, which are not on the scale. The protection of users' information is mandatory by law, so it is fundamental for all scales to include this as part of an integral evaluation of a mobile app. Likewise, the scale attaches little importance to whether an app is evidence-based or trialed in well-controlled studies. For example, a recent study that used MARS [24] to assess the quality of pain-related mobile apps showed that of the 18 apps, the 2 that had been scientifically tested were given the worst scores on the scale, and 1 of these had already been awarded a seal of quality from a public agency. It does seem that with MARS, the so-called commercial apps are better rated than those that have been scientifically tested and shown to provide valid and reliable information. This goes against the current trend in the area, which is seeking apps that have been scientifically tested and designed on the basis of evidence [25-27]. Furthermore, Salazar et al [24] showed that when MARS is used, an app developed with a highly specific objective in mind (eg, to measure pain intensity) will show lower scores (and will, therefore, be assumed to provide worse measurements) simply because of its specificity. Finally, the questions on the rating scale developed by Stoyanov et al [17] were mostly written to be answered by end users and require responses that are highly subjective or cannot be answered by a person who is not an expert in the field (eg, "Is app content correct, well written, and relevant to the goal or topic of the app?").

In addition, the preliminary data on the comprehension of the criteria showed that they can be understood by different profiles of stakeholders, as intended. However, a few of them reported having problems with some criteria, which were solved after giving additional explanations. Therefore, it is important that the information is presented with the least technical wording possible to facilitate comprehension. Nevertheless, additional studies with more participants to validate and extend the findings are warranted.

The resulting guide with this set of criteria describes the standard to follow, identifies the main categories of criteria, and provides stakeholders with a systematic approach by which they can determine the general requirements of a mobile app if it is to be considered of high quality. An app that meets these criteria is one that will provide users with the greatest security and confidence in performance and the objectives being fulfilled.

Limitations

This study has limitations that should be considered when interpreting the results. First, our search strategy was limited to papers written in English or Spanish, pain-related apps, and guidelines and standards published in specific regions and countries. We made these choices because it was what we could feasibly do, but we cannot be certain that we have included *all* the important criteria. For example, some issues could be seen as more important by developers of pain-related apps compared with developers of apps related to sexual health (eg, pain-related apps are biased toward treatment rather than diagnosis; pain-related apps may primarily be targeted at the patient, rather than health professionals or carers). We analyzed the studies on pain-related apps, and the information was combined with that from the most important markets for mHealth apps and on guidelines and standards available, as a way to complement each other and solve the potential limitations. Nevertheless, the final result of our analysis is limited in ways that we cannot completely foresee. Therefore, future studies on the validity and reliability of this set of criteria are warranted. Second, the comprehension test was conducted with a small group of 18 individuals from three groups of stakeholders. Although the number of participants was enough for a preliminary analysis, the sample is not representative. Thus, additional studies, including samples with more participants, are needed. Despite these limitations, this study provides important new information to help advance the field.

Conclusions

This set of criteria can be readily used by health care providers, engineers and developers, researchers, patients, and regulators. The data have shown them to be comprehensible and of importance for a group of stakeholders. Nevertheless, future

studies will have to empirically test the validity, reliability, and suitability of this set of criteria. Furthermore, they should be analyzed in terms of their significance to all stakeholders so that the set of criteria could also be used as a guide to the quality of the apps by all interested parties.

Acknowledgments

This work was partly supported by grants from the Spanish Ministry of Economy and Competitiveness (RTI2018-09870-B-I00; RED2018-102546-T), the European Regional Development Fund, and Obra Social de Caixabank. JM's work is supported by Fundación Grünenthal. PL benefitted from a predoctoral fellowship (2019 FI_B2 00151) cofinanced by the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya, the EU, and the European Social Fund.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full list of criteria.
[DOC File , 438 KB - mhealth_v8i3e13057_app1.doc]

References

1. OECD. Health at a Glance 2017 OECD Indicators. Paris, France: OECD; 2017.
2. World Health Organization. mHealth: New Horizons for Health through Mobile Technologies. Geneve: World Health Organization; 2011.
3. Wang Y, Sun L, Hou J. Hierarchical medical system based on big data and mobile internet: a new strategic choice in health care. JMIR Med Inform 2017 Aug 8;5(3):e22 [FREE Full text] [doi: 10.2196/medinform.6799] [Medline: 28790024]
4. Hanseth O, Bygstad B. Flexible generification: ICT standardization strategies and service innovation in health care. Eur J Inf Syst 2015;24(6):645-663. [doi: 10.1057/ejis.2015.1]
5. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. Pharm Ther 2014 May;39(5):356-364 [FREE Full text] [Medline: 24883008]
6. Ernsting C, Dombrowski SU, Oedekoven M, O Sullivan JL, Kanzler M, Kuhlmey A, et al. Using smartphones and health apps to change and manage health behaviors: a population-based survey. J Med Internet Res 2017 Apr 5;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
7. de la Vega R, Miró J. mHealth: a strategic field without a solid scientific soul. a systematic review of pain-related apps. PLoS One 2014;9(7):e101312 [FREE Full text] [doi: 10.1371/journal.pone.0101312] [Medline: 24999983]
8. Germanakos P, Mourlas C, Samaras G. A Mobile Agent Approach for Ubiquitous and Personalized eHealth Information Systems. 2005 Presented at: Proc Work ?Personalization e-Health? 10th Int Conf User Model ;67?70; 2005; Berlin.
9. Huckvale K, Prieto JT, Tilney M, Benghozi P, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015 Sep 7;13:214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]
10. Miró J. Dolor Crónico. Procedimientos de evaluación e intervención psicológica. Bilbao: Desclée de Brower; 2003.
11. Boulos MN, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014;5(3):229 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4814] [Medline: 24683442]
12. 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]
13. Tic Salut Social Foundation. 2015. The Mobility Master Plan: mHealth solutions URL: https://ticsalutsocial.cat/en/actualitat/the-mobility-master-plan-mhealth-solutions/ [accessed 2018-11-12] [WebCite Cache ID 73sMqAseU]
14. Tic Salut Social Foundation. 2015. AppSalut URL: https://ticsalutsocial.cat/en/projects/mhealth-en/appsalut-marketplace/ [accessed 2018-11-12] [WebCite Cache ID 73sLiv7ZO]
15. Tic Salut Social Foundation. 2018. Criteria Guide of Accreditation Process for Health and Wellness Apps URL: https://ticsalutsocial.cat/en/resources/mhealth-en/apps-accreditation-guidebooks/ [accessed 2018-11-12] [WebCite Cache ID 73sMDaN6Y]
16. European Commission. 2014 Apr 10. Green Paper on mobile Health ('mHealth') URL: https://ec.europa.eu/digital-single-market/en/news/green-paper-mobile-health-mhealth [accessed 2020-01-17] [WebCite Cache ID 73sKpDXVn]
17. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]

18. Brown W, Yen PY, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
19. Gaziel-Yablowitz M, Schwartz DG. A review and assessment framework for mobile-based emergency intervention apps. ACM Comput Surv 2018;51(1):1-32. [doi: 10.1145/3145846]
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
21. Wikipedia. Mobile app URL: https://en.wikipedia.org/wiki/Mobile_app [accessed 2018-11-12] [WebCite Cache ID 73sLCDxl2]
22. Statista. Global Mobile Health Market in 2017, by Region URL: https://www.statista.com/statistics/218852/global-mobile-health-market-share-forecast-by-region/ [accessed 2018-11-12] [WebCite Cache ID 73sLFZA7t]
23. Research2Guidance. mHealth Economics 2017 – Current Status and Future Trends in Mobile Health URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2020-01-17] [WebCite Cache ID 73sK8Aedl]
24. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718 [FREE Full text] [doi: 10.2196/10718] [Medline: 30361196]
25. Buijink AW, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. Evid Based Med 2013 Jun;18(3):90-92. [doi: 10.1136/eb-2012-100885] [Medline: 22923708]
26. Larson RS. A path to better-quality mHealth apps. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e10414 [FREE Full text] [doi: 10.2196/10414] [Medline: 30061091]
27. Jonassaint CR, Rao N, Sciuto A, Switzer GE, de Castro L, Kato GJ, et al. Abstract animations for the communication and assessment of pain in adults: cross-sectional feasibility study. J Med Internet Res 2018 Aug 3;20(8):e10056 [FREE Full text] [doi: 10.2196/10056] [Medline: 30076127]

Abbreviations

EU: European Union
MARS: Mobile App Rating Scale
mHealth: mobile health

Edited by G Eysenbach; submitted 07.12.18; peer-reviewed by H de Sola, N West, HT Yang, A Salazar, D Schwartz, S Wali, S Zheng; comments to author 14.04.19; revised version received 06.06.19; accepted 23.07.19; published 03.03.20.

Please cite as:
Llorens-Vernet P, Miró J
Standards for Mobile Health–Related Apps: Systematic Review and Development of a Guide
JMIR Mhealth Uhealth 2020;8(3):e13057
URL: https://mhealth.jmir.org/2020/3/e13057
doi:10.2196/13057
PMID:32130169

Original Paper

Translation of the Chinese Version of the Nomophobia Questionnaire and Its Validation Among College Students: Factor Analysis

Ye Gao[1*], BSc; Hongliang Dai[1*], PhD; Guizhi Jia[2,3*], MSc; Chunguang Liang[1], PhD; Tong Tong[1], BSc; Zhiyu Zhang[1], BSc; Ruobing Song[1], BSc; Qing Wang[1], BSc; Yue Zhu[1], BSc

[1]School of Nursing, Jinzhou Medical University, Jinzhou, China

[2]Department of Physiology, Jinzhou Medical University, Jinzhou, China

[3]Department of Physiology, Shanxi Medical University, Taiyuan, China

[*]these authors contributed equally

Corresponding Author:
Chunguang Liang, PhD
School of Nursing
Jinzhou Medical University
No 40, Section 3, Songpo Road
Jinzhou, 121001
China
Phone: 86 15941628736
Email: 63398880@qq.com

Abstract

Background: Nomophobia or phobia of no mobile phone is the fear of being without a mobile phone or being unable to contact others via a mobile phone. It is a newly emerging psychiatric disorder among mobile phone users.

Objective: There are no psychometric scales available in China for examining nomophobia, although China has become the largest mobile phone handset consumer market in the world. Therefore, this study aimed to translate the original English version of a psychometric scale into Chinese and further examine its reliability and validity among Chinese college students.

Methods: The original version of the Nomophobia Questionnaire (NMP-Q) was first translated into Chinese using the backward and forward translation procedure. An exploratory factor analysis (a principal component analysis plus varimax rotation) and a confirmatory factor analysis (CFA) were performed to examine the underlying factor structure of the translated questionnaire. The internal consistency reliability of the scale was determined by computing the Cronbach alpha coefficient, the test-retest reliability, and the corrected item-total correlation. A multivariate regression analysis was used for examining associations between nomophobia and independent variables among the college students.

Results: A total of 2000 participants were included in the study. Their ages ranged from 16 to 25 years, with 51.95% (1039/2000) being male participants. The Chinese version of NMP-Q retained 18 items. The eigenvalues, total variance explained, and scree plot jointly support a 4-factor structure of the translated questionnaire. The CFA reached the adaptive standard, and the discriminant validity of the scale was good. The Cronbach alpha coefficient of this scale was .925, and the Cronbach alpha coefficients of the subscales were .882, .843, .895, and .818. The test-retest reliability was 0.947. Corrected item-total correlation ranged from 0.539 to 0.663. The significant predictors for each of the dimensions of nomophobia and total score of the questionnaire were the average number of hours spent on a mobile phone daily and gender.

Conclusions: The Chinese version of the NMP-Q exhibited satisfactory psychometric properties.

(JMIR Mhealth Uhealth 2020;8(3):e13561) doi:10.2196/13561

KEYWORDS

nomophobia; reliability; validity; mobile phone

Introduction

Background

The mobile phone embodies the latest evolution of modern information and communication technologies [1]; in recent years, it has become increasingly widespread within the Chinese society. In fact, China has become the largest mobile phone handset consumer market in the world, and the market is expected to grow rapidly in the coming years [2]. As of December 2016, there were 695 million mobile phone subscribers with access to the internet in China, according to China's 39th Statistical Report on Internet Development [2]. Indeed, in addition to the traditional *phone* function for the purpose of communication, mobile phones nowadays are more akin to powerful portable computers, providing a number of other intellectual functions such as internet browser, camera, email services, real-time information, social networking, and personal diary. All these powerful functions of the mobile phone make our life unprecedentedly more dynamic, effective, and convenient, thereby making the use of mobile phones increasingly widespread among individuals from different countries and regions [1-3].

Despite the countless benefits associated with mobile phone use, it is severely damaging to our health, both physically and mentally. Previous studies suggest that mobile phone overuse exists among many people, which seriously interferes with their daily lives, safety, and health status. These health-related problems include blurred vision, pain in the wrists or neck, "screen dermatitis," tumors, infertility, and, in many cases, traffic accidents [4-6]. Moreover, smartphone overuse also harms mental health, leading to "smartphone dependence or addiction" [2,7,8], which is characterized by compulsive phone use, intolerance, withdrawal, and functional impairment. Other terms describing mobile phone–related mental disorders include "SMS texting addiction," "compulsive selfie-taking behavior," "sexting," "phubbing," and the recently emerged "nomophobia" [1,9-11].

Nomophobia, the contraction of no mobile phone phobia, is a recently emerging neologism to describe the anxiety and distress among mobile phone users when they are without a smartphone or mobile phone and are unable to get access to the services and real-time information it provides and feel disconnected. The characterized symptoms of nomophobia include overuse of a mobile phone and the resultant feeling of disturbance when there is a lack of network coverage, the battery is out of power, or the balance between the two is not enough. In addition, people with nomophobia would frequently check their phones for messages or missed calls, and they are repeatedly under the illusion of hearing a mobile phone ring or vibrate. All these symptoms because of the problematic use of a mobile phone would inevitably affect people's social life, work productivity, and academic performance in a negative way. To evaluate this situational phobia status affecting mobile phone users, Yildirim and Correia [12] first developed the Nomophobia Questionnaire (NMP-Q) in the United States in the English language as an effective instrument. To the best of our knowledge, since the advent of the English version of NMP-Q, it has only been translated into 4 other languages, Italian [1], Persian [13], Spanish [14], and most recently Chinese [15], with proven reliability in each version. In 2018, Jianling Ma and Chang Liu translated the English version of NMP-Q based on the cultural background and language habits of the Chinese and explored the applicability of the Chinese version of the scale among 966 college students in southern China, and the results showed that the reliability and validity were good. However, whether the NMP-Q scale can be directly used to evaluate the level of nomophobia of college students in northern China needs to be confirmed.

Objectives

The aims of this study were to translate the original NMP-Q into simplified Chinese and further confirm its reliability and validity among college students in northern China.

Methods

Instrument

NMP-Q is a 20-item scale developed by Yildirim and Correia [12] through a thorough procedure comprising qualitative and quantitative phases. NMP-Q comprises 4 factors (Factor 1: losing connectedness; Factor 2: giving up convenience; Factor 3: not being able to communicate; Factor 4: not being able to access information). A 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) is applied to each NMP-Q item, leading to a summed total score. After obtaining permission from the original authors, the Chinese NMP-Q was developed by using the original English version and the classical "backward and forward" procedure [1]. The final Chinese NMP-Q (Multimedia Appendix 1) was then administered to 2000 college students in their universities in Jinzhou, along with a general questionnaire to record the variables of age, gender, whether having a girlfriend or boyfriend, average daily hours of mobile phone use, residence, and grade.

Participants and Procedure

This was a cross-sectional study, and the participants comprised college students aged 16 to 25 years from 3 universities in Jinzhou, China. The participants who possessed a mobile phone were recruited in this study via a convenience sampling method. The research procedures complied with the ethical standards of the institutional research committee, as well as the 1964 Helsinki declaration and its later amendments. All these students had been given written informed consent.

Statistical Analysis

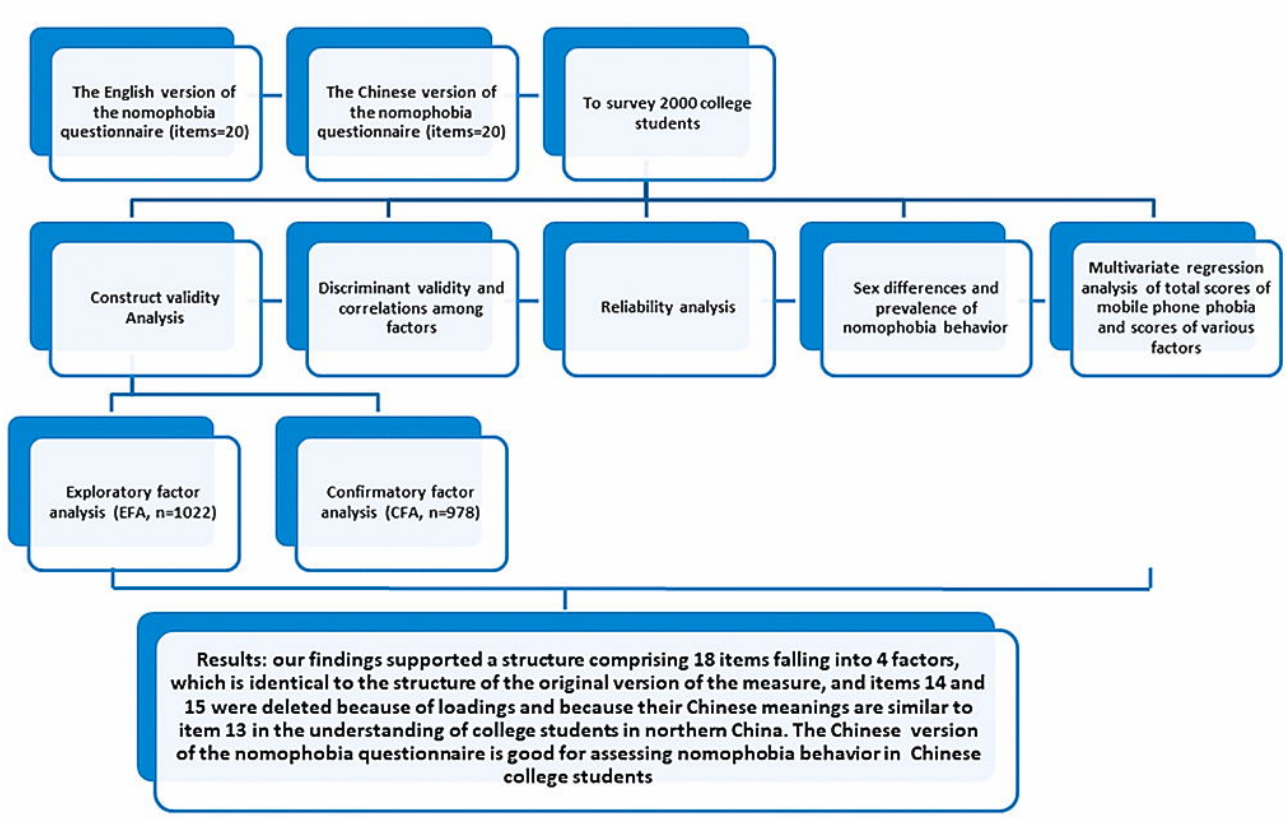
Statistical analysis was performed using SPSS 22.0 and AMOS 21.0 (IBM Corporation). Continuous data were presented as mean (SD), whereas categorical data were expressed as percentages. Skewness and kurtosis were computed for each item, and the data were considered normally distributed when the values ranged from -2 to +2 [1]. To investigate the underlying factor structure of the translated questionnaire, an exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA) were performed. The sample of 2000 cases was randomly divided into 2 groups, one (n=1022) for EFA and the other (n=978) for CFA. In EFA (n=1022), a principal component

analysis (PCA) with varimax rotation was performed on the 20 items of the questionnaire. Varimax rotation is the most commonly used orthogonal technique that minimizes factor complexity with a maximized variance of factor loading. The sampling adequacy for the factorability was assessed using the Kaiser-Meyer-Olkin (KMO) [16] measurement and Bartlett test [17] of sphericity. Only when the Bartlett test of sphericity was significant (P<.05) and the KMO was >0.60, the dataset was considered appropriate for PCA. The factors were extracted based on the comprehensive consideration of eigenvalues, explained total variance, and a visual inspection of the scree plot. AMOS (IBM Corporation) was used to perform the CFAs of NMP-Q, analyzing the fit of models and its respective parameter estimates.

Discriminant validity is the total score of the NMP-Q scale, which was ranked from high to low; the top 27% of the scores were grouped into the high-score group, the bottom 27% of the scores were grouped into the low-score group, and the scores of each item in the 2 groups were analyzed by using a two-tailed independent samples t test. If the scores of each item in the 2 groups reached the significance level (P<.05), discriminant validity was considered good.

Reliability for internal consistency of the scale was examined by calculating the Cronbach alpha coefficient, the test-retest reliability, and the corrected item-total correlation. The minimum acceptable Cronbach alpha coefficient was set at .7 [18]. The corrected item-total correlation, representing the correlation between each item and the sum of the other items in a scale, was performed using the standard of 0.3 for inclusion [18,19]. A multivariate regression analysis was performed to explore the underlying independent variables of nomophobia. The flow of the study is presented in Figure 1.

Figure 1. Diagram showing the flow of the study.



Results

Descriptive Statistics

This investigation was conducted among 2000 college students, with ages ranging from 16 to 25 years. A total of 51.95% (1039/2000) of participants were males. Regarding daily mobile phone use, 5.70% (114/2000) of participants usually spent less than 2 hours on their mobile phones, 25.35% (507/2000) of participants spent between 2 and 4 hours, 42.40% (848/2000) of participants spent between 4 and 6 hours, 13.85% (277/2000) of participants spent between 6 and 8 hours, 6.65% (133/2000) of participants spent between 8 and 10 hours, and 6.05% (121/2000) of participants spent more than 10 hours.

The mean (SD) score of each item of the Chinese NMP-Q is shown in Table 1. These data were normally distributed according to the skewness and kurtosis figures.

Table 1. Mean (SD) scores with skewness and kurtosis figures (N=2000).

Item number	Mean (SD)	Skewness	Kurtosis
1	3.61 (1.785)	0.193	-1.025
2	3.57 (1.787)	0.215	-1.078
3	3.10 (1.636)	0.539	-0.616
4	4.04 (1.808)	-0.178	-1.127
5	4.28 (1.907)	-0.311	-1.146
6	4.14 (1.882)	-0.214	-1.169
7	4.11 (1.958)	-0.158	-1.284
8	4.24 (1.850)	-0.249	-1.125
9	4.47 (1.737)	-0.470	-0.785
10	4.74 (1.761)	-0.653	-0.617
11	4.75 (1.764)	-0.660	-0.597
12	4.62 (1.763)	-0.563	-0.723
13	4.17 (1.791)	-0.178	-1.027
14	3.96 (1.785)	-0.061	-1.076
15	4.01 (1.794)	-0.102	-1.055
16	3.63 (1.772)	0.192	-0.992
17	3.24 (1.711)	0.440	-0.757
18	3.51 (1.778)	0.226	-1.022
19	2.87 (1.649)	0.713	-0.332
20	3.45 (1.828)	0.248	-1.045

Construct Validity Analysis

Before commencing an EFA, the factorability of the matrix of a sample (n=1022) was first examined. The Bartlett test [17] of sphericity was significant (χ^2_{190}=12,413.0; P<.001), and the KMO index [16] was 0.934, which is greater than the minimum acceptable value of 0.6. Therefore, the matrix is not an identity matrix and is appropriate for factor extraction.

It was shown that the vast majority of the items (16/20) were loaded on a single factor, and the loadings on other factors were much lower. The loading of items 4, 14, 15, and 20 was greater than 0.4 on 2 different factors [20]. After a discussion with experts, items 14 and 15 were deleted as they were similar to item 13 in Chinese. For college students in northern China, items 4 and 20 were reserved because of their influence and contribution rate on the scale structure. A re-exploratory factor analysis was performed after deletion; the Bartlett test [17] of sphericity was significant (χ^2_{153}=10,609.0; P<.001), and the KMO index [16] was 0.928. The results are shown in Table 2.

Table 2. Factor loadings of the Nomophobia Questionnaire (N=1022; salient factor loadings are indicated in italics).

Item number	Factor 1	Factor 2	Factor 3	Factor 4
17	*0.846*	0.123	0.133	0.235
19	*0.834*	0.051	0.092	0.163
18	*0.800*	0.255	0.109	0.154
16	*0.746*	0.257	0.211	0.136
20	*0.584*	0.452	0.108	0.079
6	0.163	*0.800*	0.104	0.248
5	0.121	*0.753*	0.187	0.314
9	0.233	*0.679*	0.287	0.118
7	0.239	*0.655*	0.180	0.179
8	0.189	*0.564*	0.372	0.209
11	0.086	0.193	*0.889*	0.104
12	0.131	0.172	*0.850*	0.116
10	0.086	0.307	*0.817*	0.110
13	0.291	0.116	*0.673*	0.278
2	0.139	0.292	0.160	*0.770*
3	0.385	0.078	0.151	*0.754*
4	0.070	0.481	0.153	*0.646*
1	0.227	0.342	0.181	*0.586*

The first PCA was run to determine the likely number of factors. As a result, 4 factors that explained a total of 68.93% of the variance had initial eigenvalues >1 each. The 4-factor structure was further confirmed by the scree plot, as the descending tendency became weak after the fourth point. After varimax rotation, these 4 extracted factors explained 43.39%, 11.24%, 8.66%, and 5.65% of the variance. The scree plot is shown in Figure 2. A CFA was performed on the sample (n=978). In the model fitness index, the chi-square degree of freedom was 4.967, goodness-of-fit index was 0.933, adjusted goodness-of-fit index was 0.909, parsimonious goodness-of-fit index was 0.687, incremental fit index was 0.952, Tucker Lewis index was 0.942, comparative fit index was 0.952, root mean square error of approximation was 0.064, and standardized root mean residual was 0.049. The CFA results are shown in Figure 3.

Figure 2. Screen plot of exploratory factor analysis for Chinese version of the Nomophobia Questionnaire.

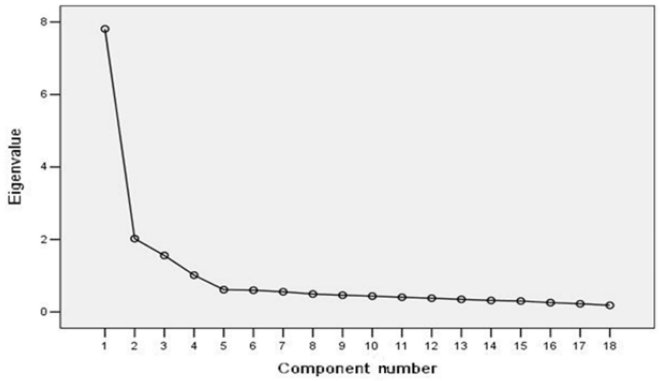
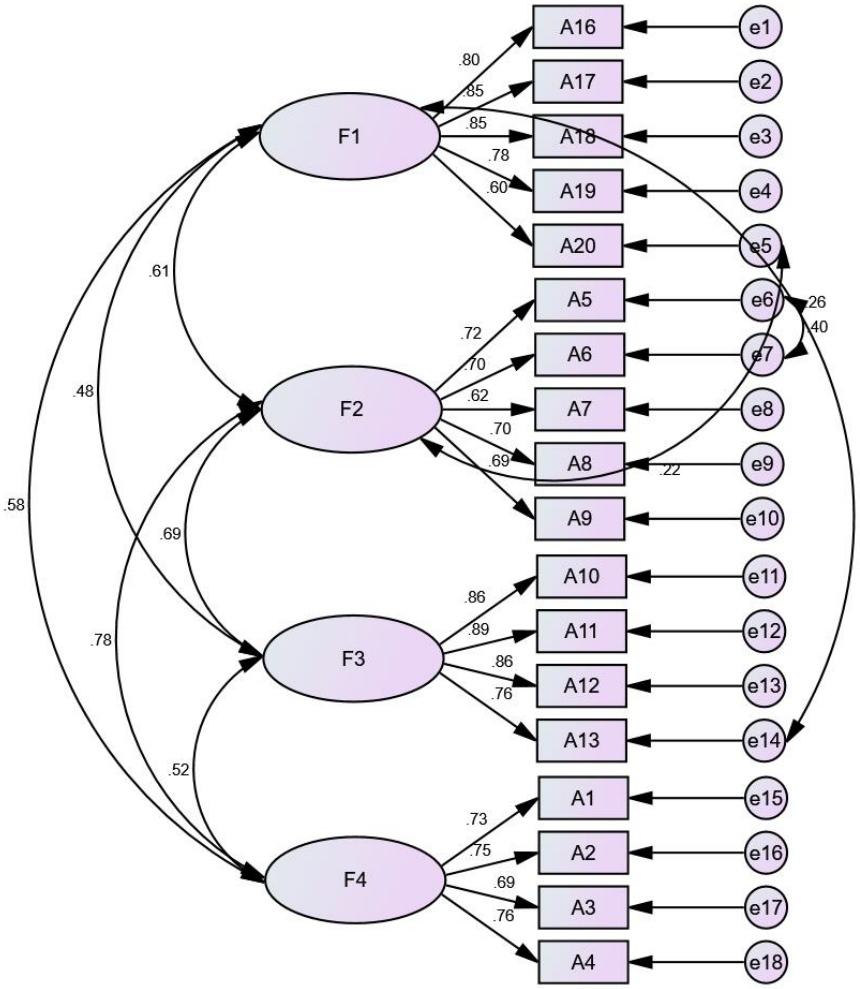


Figure 3. Standardized four-factor structural model of the Nomophobia Questionnaire (n=978). F1 (losing connectedness, five items), F2 (giving up convenience, five items), F3 (not being able to communicate, four items), F4 (not being able to access information, four items).



Discriminant Validity and Correlations Among Factors

The total score of the NMP-Q scale was ranked from high to low; the top 27% of the scores were grouped into the high-score group, and the bottom 27% of the scores were grouped into the low-score group. In this study, critical value scores were 83 and 58, respectively, and the score of each item in the 2 groups was analyzed by using a 2-tailed independent samples *t* test. The results showed that the score of each item in the 2 groups reached the level of significance ($P<.05$). The results are shown in Table 3. The results of correlations among factors in the Chinese version of NMP-Q are shown in Table 4.

Table 3. Score comparison between high-score and low-score groups (N=2000).

Item	Low-score group (n=553), mean (SD)	High-score group (n=567), mean (SD)	*t* test (*df*)	*P* value
1	2.21 (1.286)	5.07 (1.444)	-35.077 (1108.951)	<.001
2	2.22 (1.319)	5.00 (1.480)	-33.217 (1109.057)	<.001
3	1.90 (1.024)	4.35 (1.596)	-30.634 (967.443)	<.001
4	2.52 (1.514)	5.51 (1.165)	-36.917 (1036.226)	<.001
5	2.52 (1.536)	5.86 (1.099)	-41.765 (998.128)	<.001
6	2.45 (1.479)	5.64 (1.274)	-38.591 (1085.682)	<.001
7	2.58 (1.584)	5.63 (1.351)	-34.609 (1081.911)	<.001
8	2.62 (1.521)	5.71 (1.115)	-38.609 (1011.067)	<.001
9	2.98 (1.673)	5.76 (1.044)	-33.304 (921.158)	<.001
10	3.21 (1.802)	6.00 (0.929)	-32.446 (821.716)	<.001
11	3.29 (1.866)	5.98 (0.959)	-30.136 (819.922)	<.001
12	3.18 (1.759)	5.85 (1.093)	-30.476 (918.955)	<.001
13	2.61 (1.474)	5.57 (1.216)	-36.585 (1068.760)	<.001
16	2.20 (1.319)	5.15 (1.412)	-36.109 (1115.903)	<.001
17	1.94 (1.125)	4.65 (1.525)	-33.970 (1041.411)	<.001
18	2.05 (1.213)	4.97 (1.415)	-37.104 (1100.063)	<.001
19	1.80 (1.000)	4.02 (1.743)	-26.223 (906.649)	<.001
20	2.06 (1.245)	4.86 (1.629)	-32.421 (1057.658)	<.001

Table 4. Correlations among factors in the Chinese version of the Nomophobia Questionnaire (N=2000).

Factor	Factor 1	Factor 2	Factor 3
Factor 2	0.556[a]	—[b]	—
Factor 3	0.443[a]	0.569[a]	—
Factor 4	0.542[a]	0.663[a]	0.465[a]

[a]Significant correlation at the .01 level (two-sided).

[b]Not available.

Reliability Analysis

Reliability analysis results showed that the Chinese version of NMP-Q has ideal internal consistency, with the overall Cronbach alpha coefficient being .925 and Cronbach alpha coefficients for the 4 factors being .882, .843, .895, and .818, all greater than the minimum acceptable value of .7. The change in the Cronbach alpha value when a given item is excluded from the questionnaire is listed in Table 5. As seen in Table 5, the deletion of each item each time exclusively led to a decrease in the Cronbach alpha coefficient of the questionnaire (.925). In addition, the corrected item-total correlation of the items ranged from 0.539 to 0.663, that is, all greater than 0.3. As such, 18 items are integral to the questionnaire. After 2 weeks, 30 students were randomly selected for retesting; the test-retest reliability was 0.947.

Table 5. Cronbach alpha coefficient if the item was deleted and corrected item-total correlation (N=2000).

Item	Cronbach alpha if the item was deleted	Corrected item-total correlation
1	.921	0.600
2	.922	0.585
3	.922	0.577
4	.921	0.632
5	.920	0.663
6	.920	0.649
7	.922	0.586
8	.921	0.637
9	.921	0.620
10	.921	0.617
11	.922	0.588
12	.922	0.592
13	.921	0.636
16	.920	0.652
17	.921	0.622
18	.921	0.634
19	.923	0.539
20	.922	0.596

Sex Differences and Prevalence of Nomophobic Behavior

The independent *t* test found that females (mean 75.61, SD 20.32) had a higher total nomophobia score than males (mean 65.82, SD 21.23; t_{2000}= -10.52; *P*<.001). To evaluate the prevalence of nomophobic behavior in Chinese college students, a total nomophobia score was calculated for each participant and transformed into a Z-score in the sample (N=2000). A Z-score lower than -1 was regarded as demonstrating the absence of nomophobia. A Z-score between -1 and the mean was seen as demonstrating a low level of nomophobia. A Z-score above the mean but lower than 1 was taken as demonstrating a mild level of nomophobia. Finally, a Z-score greater than 1 but lower than 2 was recognized as demonstrating a severe level of nomophobia, and a Z-score above 2 was classified as a very severe level of nomophobia. Specifically, there were 15.85% (317/2000), 31.75% (635/2000), 36.95% (739/2000), 13.20% (264/2000), and 2.25% (45/2000) of participants in the without, low, mild, severe, and very severe levels of nomophobia groups, respectively.

Multivariate Regression Analysis of the Total Scores of Nomophobia and Scores of Various Factors

On the basis of the multivariate regression analysis results shown in Table 6, the NMP-Q total score correlated with daily hours of mobile phone use, gender, and residence, although no statistically significant associations were found with age, grade, and whether or not having a boyfriend or girlfriend.

Identical patterns were found for the total score of the subscale, “not being able to access information.” The total score of the subscale, “giving up convenience,” showed correlations with daily hours of mobile phone use, gender, and grade. “Not being able to communicate” correlated with daily hours of mobile phone use, age, gender, grade, and residence. “Losing connectedness” correlated with daily hours of mobile phone use, gender, grade, and residence.

In conclusion, the daily hours of mobile phone use and gender correlated with all the 4 subscales of the questionnaire, whereas age, grade, residence, and whether or not having a boyfriend or girlfriend only associated with certain subscales. Age had no association with the 4 subscales.

Table 6. A multivariate regression analysis on the impact of variables on total scores of the scale and subscales (N=2000).

Model	B	SD	Beta	*t* test (*df*)	*P* value
Overall					
Constant	30.910	8.597	—[a]	3.595 (6)	<.001
Number of hours of mobile phone use	3.497	0.383	0.195	9.131 (6)	<.001
Age (years)	0.726	0.442	0.042	1.643 (6)	.10
Gender	8.927	0.917	0.209	9.740 (6)	<.001
Grade	0.520	0.866	0.016	0.600 (6)	.55
Residence	2.545	0.913	0.060	2.787 (6)	.005
Having a boyfriend or girlfriend or not	-1.589	1.013	-0.034	-1.569 (6)	.12
Not being able to access information					
Constant	5.279	2.311	—	2.285 (6)	.02
Number of hours of mobile phone use	0.683	0.103	0.144	6.632 (6)	<.001
Age (years)	0.144	0.119	0.032	1.210 (6)	.23
Gender	2.133	0.246	0.189	8.659 (6)	<.001
Grade	0.273	0.233	0.031	1.171 (6)	.24
Residence	0.523	0.245	0.046	2.132 (6)	.03
Having a boyfriend or girlfriend or not	-0.063	0.272	-0.005	-0.233 (6)	.82
Giving up convenience					
Constant	9.967	2.959	—	3.369 (6)	.001
Number of hours of mobile phone use	1.273	0.132	0.208	9.657 (6)	<.001
Age (years)	0.225	0.152	0.038	1.482 (6)	.14
Gender	2.852	0.315	0.195	9.040 (6)	<.001
Grade	-0.850	0.298	-0.074	-2.850 (6)	.004
Residence	0.509	0.314	0.035	1.621 (6)	.11
Having a boyfriend or girlfriend or not	-0.449	0.349	-0.028	-1.286 (6)	.20
Not being able to communicate					
Constant	9.000	2.552	—	3.527 (6)	<.001
Number of hours of mobile phone use	0.463	0.114	0.089	4.069 (6)	<.001
Age (years)	0.305	0.131	0.062	2.329 (6)	.02
Gender	2.169	0.272	0.176	7.974 (6)	<.001
Grade	-0.882	0.257	-0.091	-3.430 (6)	.001
Residence	0.671	0.271	0.054	2.474 (6)	.01
Having a boyfriend or girlfriend or not	-0.568	0.301	-0.042	-1.888 (6)	.06
Losing connectedness					
Constant	6.663	2.905	—	2.294 (6)	.02
Number of hours of mobile phone use	1.079	0.129	0.179	8.339 (6)	<.001
Age (years)	0.052	0.149	0.009	0.347 (6)	.73
Gender	1.774	0.310	0.123	5.727 (6)	<.001
Grade	1.979	0.293	0.175	6.761 (6)	<.001
Residence	0.841	0.309	0.058	2.727 (6)	.006
Having a boyfriend or girlfriend or not	-0.510	0.342	-0.032	-1.489 (6)	.14

[a]Not available.

Discussion

Principal Findings

This study translated NMP-Q, developed by Yildirim and Correia [12], into Chinese and tested its psychometric properties among a large number of Chinese college students to investigate whether the Chinese NMP-Q can be applied to college students in China. Our findings supported a structure comprising 18 items falling into 4 factors, which is identical to the structure of the original version of the measure; items 14 and 15 were deleted because of double loadings and because their Chinese meanings are similar to item 13 in the understanding of college students in northern China. Construct validity, discriminant validity, and reliability analysis confirmed that the scale is reliable and valid for assessing nomophobic behavior in Chinese college students.

To the best of our knowledge, our results differ from those of Jianling Ma and Chang Liu in China [15]. One possible reason might be the cultural, territorial, and sample differences. China is a vast country divided into north and south regions, and people's personalities are different; the southern people are more sensitive than the northern people. The sample of our survey was from northern China, whereas the sample that Jianling Ma and Chang Liu surveyed was from southern China. Therefore, the understanding of items 13, 14, and 15 may be different. The other reason is the sample size (2000 students instead of 966). Therefore, our results may be more appropriate for college students in northern China. The original scale has already been validated in Spain [14], Iran [13], and Italy [1], and in these cases, the content and structural validity of the scale were supported. The Spanish version of NMP-Q showed 4 factors consistent with the original scale. However, in the Persian version, although the 4-factor model was also supported, items 9 and 14 were deleted because of low factor loadings [13]. The Italian version of NMP-Q showed 3 factors [1]. One possible reason might be the cultural and sample differences: although we selected college students, the Persian version used teenagers aged between 13 and 19 years. The Italian version surveyed 403 people, and we surveyed 2000 people, and the people we surveyed were younger than those in the Italian sample (mean 19.07, SD 1.25 years vs mean 27.91, SD 8.63 years, respectively). Therefore, the results are different.

The total scores, both of the scales and all subscales, correlated with the daily hours of mobile phone use and gender, which confirmed the validation of the Chinese version of NMP-Q. Another interesting finding of this study is that gender had a significant influence on nomophobia; specifically, female students were more negatively affected by being without a mobile phone compared with their male counterparts. Our present data are in line with previous studies showing that females were at greater risk of mobile phone addiction [21-23], which may be related to women's stronger desire for social relationships [21]. In fact, the association between gender and problematic use of mobile phones is quite complex, and, to date, the investigation regarding the effect of gender on nomophobia has produced conflicting results [1,21,24,25]. The discrepancy among studies, including ours, is difficult to understand and might be related to the difference in ethnicity and subgroup of the population included in different studies. In addition, the differential interpretation of the item content and factor structures between genders because of different brain structures might also have an impact on the results [13,26]; this issue warrants further exploration.

Several limitations should be taken into account when interpreting the findings in this study. First, the sample was conveniently obtained; therefore, there is a lack of representativeness. Furthermore, only college students were included in the study. Hence, our results should be generalized with caution. Second, bias was inevitable because of the self-reporting nature of this investigation. Third, data regarding other psychiatric disorders, such as depression, anxiety, and obsessive-compulsive disorders, have not been taken into account, and these may become underlying confounding factors. Besides, because of a dearth of research on this topic, it is difficult to have a proper discussion involving a comparison with the findings of previous studies. Nevertheless, this concurrently highlights the novelty, fun, and readability of this study.

Conclusions

Nomophobia is an emerging psychological disorder related to problematic mobile phone use and is increasingly studied by health care researchers. Research on this condition might have greater significance for China, as it has become the largest mobile phone handset consumer market in the world. The Chinese version of this instrument, supporting a 4-factor structure, turned out to be reliable; therefore, it can be employed for investigating nomophobia in the Chinese society. Future research should be encouraged to examine the psychometric properties of this translated NMP-Q across different groups in China, to determine underlying comorbidities, and to explore the relationship of nomophobia with other technopathy. In addition, the underlying predictors of nomophobia should be further identified.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Chinese version of the Nomophobia Questionnaire.
[PDF File (Adobe PDF File), 112 KB - mhealth_v8i3e13561_app1.pdf]

References

1. Adawi M, Bragazzi NL, Argumosa-Villar L, Boada-Grau J, Vigil-Colet A, Yildirim C, et al. Translation and validation of the nomophobia questionnaire in the Italian language: exploratory factor analysis. JMIR Mhealth Uhealth 2018 Jan 22;6(1):e24 [FREE Full text] [doi: 10.2196/mhealth.9186] [Medline: 29358161]
2. Chen B, Liu F, Ding S, Ying X, Wang L, Wen Y. Gender differences in factors associated with smartphone addiction: a cross-sectional study among medical college students. BMC Psychiatry 2017 Oct 10;17(1):341 [FREE Full text] [doi: 10.1186/s12888-017-1503-z] [Medline: 29017482]
3. Prasad M, Patthi B, Singla A, Gupta R, Saha S, Kumar JK, et al. Nomophobia: a cross-sectional study to assess mobile phone usage among dental students. J Clin Diagn Res 2017 Feb;11(2):ZC34-ZC39 [FREE Full text] [doi: 10.7860/JCDR/2017/20858.9341] [Medline: 28384977]
4. Kwon M, Lee JY, Won WY, Park JW, Min JA, Hahn C, et al. Development and validation of a smartphone addiction scale (SAS). PLoS One 2013;8(2):e56936 [FREE Full text] [doi: 10.1371/journal.pone.0056936] [Medline: 23468893]
5. Kesari KK, Siddiqui MH, Meena R, Verma HN, Kumar S. Cell phone radiation exposure on brain and associated biological systems. Indian J Exp Biol 2013 Mar;51(3):187-200. [Medline: 23678539]
6. Makker K, Varghese A, Desai NR, Mouradi R, Agarwal A. Cell phones: modern man's nemesis? Reprod Biomed Online 2009 Jan;18(1):148-157. [doi: 10.1016/s1472-6483(10)60437-3] [Medline: 19146782]
7. Albursan IS, Al Qudah MF, Dutton E, Hassan EM, Bakhiet SF, Alfnan AA, et al. National, sex and academic discipline difference in smartphone addiction: a study of students in Jordan, Saudi Arabia, Yemen and Sudan. Community Ment Health J 2019 Jul;55(5):825-830. [doi: 10.1007/s10597-019-00368-x] [Medline: 30661142]
8. Lee EJ, Kim HS. Gender differences in smartphone addiction behaviors associated with parent-child bonding, parent-child communication, and parental mediation among Korean elementary school students. J Addict Nurs 2018;29(4):244-254. [doi: 10.1097/JAN.0000000000000254] [Medline: 30507820]
9. Bianchi D, Morelli M, Nappa MR, Baiocco R, Chirumbolo A. A bad romance: sexting motivations and teen dating violence. J Interpers Violence 2018 Dec 10:886260518817037. [doi: 10.1177/0886260518817037] [Medline: 30537886]
10. Franchina V, Abeele MV, van Rooij AJ, Lo Coco G, de Marez L. Fear of missing out as a predictor of problematic social media use and phubbing behavior among flemish adolescents. Int J Environ Res Public Health 2018 Oct 22;15(10):pii: E2319 [FREE Full text] [doi: 10.3390/ijerph15102319] [Medline: 30360407]
11. Aguilera-Manrique G, Márquez-Hernández VV, Alcaraz-Córdoba T, Granados-Gámez G, Gutiérrez-Puertas V, Gutiérrez-Puertas L. The relationship between nomophobia and the distraction associated with smartphone use among nursing students in their clinical practicum. PLoS One 2018;13(8):e0202953 [FREE Full text] [doi: 10.1371/journal.pone.0202953] [Medline: 30148870]
12. Yildirim C, Correia AP. Exploring the dimensions of nomophobia: development and validation of a self-reported questionnaire. Comput Human Behav 2015;49:130-137. [doi: 10.1016/j.chb.2015.02.059]
13. Lin CY, Griffiths MD, Pakpour AH. Psychometric evaluation of Persian Nomophobia Questionnaire: Differential item functioning and measurement invariance across gender. J Behav Addict 2018 Mar 1;7(1):100-108 [FREE Full text] [doi: 10.1556/2006.7.2018.11] [Medline: 29444607]
14. González-Cabrera J, León-Mejía A, Pérez-Sancho C, Calvete E. Adaptation of the Nomophobia Questionnaire (NMP-Q) to Spanish in a sample of adolescents. Actas Esp Psiquiatr 2017 Jul;45(4):137-144 [FREE Full text] [Medline: 28745386]
15. Ma J, Liu C. Evaluation of the factor structure of the Chinese version of the nomophobia questionnaire. Curr Psychol 2018:1-7. [doi: 10.1007/s12144-018-0071-9]
16. Kaiser HF. An index of factorial simplicity. Psychometrika 1974;39(1):31-36. [doi: 10.1007/BF02291575]
17. Bartlett MS. A note on the multiplying factors for various $\chi 2$ approximations. J R Stat Soc Series B Stat Methodol 1954 Jul;16(2):296-298. [doi: 10.1111/j.2517-6161.1954.tb00174.x]
18. Chang Q, Sha F, Chan CH, Yip PS. Validation of an abbreviated version of the Lubben Social Network Scale ('LSNS-6') and its associations with suicidality among older adults in China. PLoS One 2018;13(8):e0201612 [FREE Full text] [doi: 10.1371/journal.pone.0201612] [Medline: 30071067]
19. Ferketich S. Focus on psychometrics. Aspects of item analysis. Res Nurs Health 1991 Apr;14(2):165-168. [doi: 10.1002/nur.4770140211] [Medline: 2047538]
20. Boone KB, Ponton MO, Gorsuch RL, Gonzalez JJ, Miller BL. Factor Analysis of Four Measures of Prefrontal Lobe Functioning. Arch Clin Neuropsychol 1998 Oct;13(7):585-595. [doi: 10.1093/arclin/13.7.585] [Medline: 14590619]
21. Lee H, Kim JW, Choi TY. Risk factors for smartphone addiction in Korean adolescents: smartphone use patterns. J Korean Med Sci 2017 Oct;32(10):1674-1679 [FREE Full text] [doi: 10.3346/jkms.2017.32.10.1674] [Medline: 28875613]
22. Choi SW, Kim DJ, Choi JS, Ahn H, Choi EJ, Song WY, et al. Comparison of risk and protective factors associated with smartphone addiction and Internet addiction. J Behav Addict 2015 Dec;4(4):308-314 [FREE Full text] [doi: 10.1556/2006.4.2015.043] [Medline: 26690626]
23. Lee H, Seo MJ, Choi TY. The effect of home-based daily journal writing in Korean adolescents with smartphone addiction. J Korean Med Sci 2016 May;31(5):764-769 [FREE Full text] [doi: 10.3346/jkms.2016.31.5.764] [Medline: 27134499]

24. Basu S, Garg S, Singh MM, Kohli C. Addiction-like behavior associated with mobile phone usage among medical students in Delhi. Indian J Psychol Med 2018;40(5):446-451 [FREE Full text] [doi: 10.4103/IJPSYM.IJPSYM_59_18] [Medline: 30275620]
25. Boumosleh JM, Jaalouk D. Depression, anxiety, and smartphone addiction in university students- a cross sectional study. PLoS One 2017;12(8):e0182239 [FREE Full text] [doi: 10.1371/journal.pone.0182239] [Medline: 28777828]
26. Ingalhalikar M, Smith A, Parker D, Satterthwaite TD, Elliott MA, Ruparel K, et al. Sex differences in the structural connectome of the human brain. Proc Natl Acad Sci U S A 2014 Jan 14;111(2):823-828 [FREE Full text] [doi: 10.1073/pnas.1316909110] [Medline: 24297904]

Abbreviations

CFA: confirmatory factor analysis
EFA: exploratory factor analysis
KMO: Kaiser-Meyer-Olkin
NMP-Q: Nomophobia Questionnaire
PCA: principal component analysis

Edited by G Eysenbach; submitted 30.01.19; peer-reviewed by N Bragazzi, L Gutierrez-Puertas; comments to author 02.10.19; revised version received 14.11.19; accepted 16.12.19; published 13.03.20.

Please cite as:
Gao Y, Dai H, Jia G, Liang C, Tong T, Zhang Z, Song R, Wang Q, Zhu Y
Translation of the Chinese Version of the Nomophobia Questionnaire and Its Validation Among College Students: Factor Analysis
JMIR Mhealth Uhealth 2020;8(3):e13561
URL: http://mhealth.jmir.org/2020/3/e13561/
doi:10.2196/13561
PMID:32167480

Original Paper

The German Version of the Mobile App Rating Scale (MARS-G): Development and Validation Study

Eva-Maria Messner[1], MA; Yannik Terhorst[1], MSc; Antonia Barke[2], DPhil; Harald Baumeister[1], PhD, Prof Dr; Stoyan Stoyanov[3], MRES; Leanne Hides[4], PhD, Prof Dr; David Kavanagh[5], PhD, Prof Dr; Rüdiger Pryss[6], PhD, Prof Dr; Lasse Sander[7], PhD; Thomas Probst[8], PhD, Prof Dr

[1]Clinical Psychology and Psychotherapy, Ulm University, Ulm, Germany
[2]Clinical and Biological Psychology, Catholic University Eichstaett-Ingolstadt, Eichstaett-Ingolstadt, Germany
[3]Creative Industries Faculty, Queensland University of Technology, Brisbane, Australia
[4]Faculty of Health and Behavioural Sciences, University of Queensland, Brisbane, Australia
[5]Psychology and Counselling, Queensland University of Technology, Brisbane, Australia
[6]Institute of Databases and Information Systems, Ulm University, Ulm, Germany
[7]Department of Rehabilitation Psychology and Psychotherapy, University of Freiburg, Freiburg, Germany
[8]Department for Psychotherapy and Biopsychosocial Health, Danube University Krems, Krems, Austria

Corresponding Author:
Eva-Maria Messner, MA
Clinical Psychology and Psychotherapy
Ulm University
Albert-Einstein-Allee 47
Ulm
Germany
Phone: 49 07315032802
Email: eva-maria.messner@uni-ulm.de

Abstract

Background: The number of mobile health apps (MHAs), which are developed to promote healthy behaviors, prevent disease onset, manage and cure diseases, or assist with rehabilitation measures, has exploded. App store star ratings and descriptions usually provide insufficient or even false information about app quality, although they are popular among end users. A rigorous systematic approach to establish and evaluate the quality of MHAs is urgently needed. The Mobile App Rating Scale (MARS) is an assessment tool that facilitates the objective and systematic evaluation of the quality of MHAs. However, a German MARS is currently not available.

Objective: The aim of this study was to translate and validate a German version of the MARS (MARS-G).

Methods: The original 19-item MARS was forward and backward translated twice, and the MARS-G was created. App description items were extended, and 104 MHAs were rated twice by eight independent bilingual researchers, using the MARS-G and MARS. The internal consistency, validity, and reliability of both scales were assessed. Mokken scale analysis was used to investigate the scalability of the overall scores.

Results: The retranslated scale showed excellent alignment with the original MARS. Additionally, the properties of the MARS-G were comparable to those of the original MARS. The internal consistency was good for all subscales (ie, omega ranged from 0.72 to 0.91). The correlation coefficients (r) between the dimensions of the MARS-G and MARS ranged from 0.93 to 0.98. The scalability of the MARS (H=0.50) and MARS-G (H=0.48) were good.

Conclusions: The MARS-G is a reliable and valid tool for experts and stakeholders to assess the quality of health apps in German-speaking populations. The overall score is a reliable quality indicator. However, further studies are needed to assess the factorial structure of the MARS and MARS-G.

(JMIR Mhealth Uhealth 2020;8(3):e14479) doi:10.2196/14479

KEYWORDS

mHealth; Mobile App Rating Scale; mobile app; assessment; rating; scale development

Introduction

Mobile phones are an integral part of modern life. In Europe, 67% of the population owns smartphones, and the number of smartphone users is rising worldwide [1]. It has been reported that 30% of Germans have 11 to 20 apps installed on their smartphones [2]. The use of mobile apps to improve mental health and well-being is becoming increasingly common, with roughly 29% of Germans currently using at least one health app [3].

Globally, between 95 million and 130 million people speak German, making it the 11th most spoken language worldwide [4,5]. Elderly individuals and people with basic education are commonly monolingual in Germany [6]. Yet, these populations have a high need for assistance in developing and maintaining health behaviors and could benefit from the use of mobile health apps (MHAs). A German MHA rating scale could help researchers and health care providers assess the quality of health apps quickly and reliably in their mother tongue. Furthermore, it would be easy to rate a German app with a German scale.

MHAs offer unique and diverse possibilities for health promotion. They allow ecological momentary assessments [7,8] and interventions [8,9]. Additionally, they can be used irrespectively of geographical, financial, and social conditions; can simultaneously target nonclinical and clinical populations; and have the capacity to provide diverse health-management strategies in an ecological setting [10]. Moreover, they support individuals, including those from high-need high-cost populations (eg, those with chronic or lifestyle diseases), in managing their health [9]; reduce help-seeking barriers; and offer a wide range of engagement options [10].

Despite the recent proliferation of MHAs [11], there are no universally accepted criteria for measuring and reporting their quality [8,12]. Therefore, it is necessary to support researchers, users, and health care providers (eg, physicians, psychotherapists, and physiotherapists) in selecting high-quality MHAs. Safe and reliable use of MHAs requires evidence of efficacy and quality, information about data protection, information about routines for emergencies (eg, self-harm and adverse effects), and overall consideration of associated risks [10].

Boudreaux and colleagues [13] suggested the following seven strategies to evaluate MHA quality: (1) Review the scientific literature; (2) Search app clearinghouse websites; (3) Search app stores; (4) Review app descriptions, user ratings, and user reviews; (5) Conduct a social media query within professional and, if available, patient networks; (6) Pilot the apps; and (7) Elicit feedback from users. This process might be too demanding for health care providers and end users when making treatment choices. A standardized and reliable quality assessment tool could facilitate this process.

Several MHA evaluation scales exist to date. The American Psychological Association released an app evaluation model comprising 33 items across the following five scales: background information, risk/privacy and security, evidence, ease of use, and interoperability [12]. The main aim of this model is to assess the likelihood of harm [9]. However, the validity and reliability assessment of this rating instrument has not yet been reported, and there is no agreement regarding its application [14].

Baumel and colleagues [15] developed the Evaluation Tool for Mobile and Web-Based eHealth Interventions (ENLIGHT), according to a comprehensive systematic review of relevant criteria. The tool allows the evaluation of app quality in terms of seven dimensions (usability, visual design, user engagement, content, therapeutic persuasiveness, therapeutic alliance, and general subjective evaluation) with 28 items. ENLIGHT also provides a checklist to assess credibility, evidence base, privacy explanation, and basic security.

The Mobile App Rating Scale (MARS) [16] is the most commonly used app evaluation tool that allows electronic health experts to rate MHAs. It includes 19 items comprising four subscales on objective MHA characteristics (engagement, functionality, esthetics, and information quality) and a further 10 items comprising two subscales on subjective characteristics (subjective app quality and perceived impact). The subscale and overall scores indicate the quality of MHAs. The MARS has been used to scientifically assess app quality in the following fields: weight management, physical activity, heart failure, diet in children and adolescents, medication adherence, mindfulness, back pain, chronic pain, smoking cessation, and depression [8,17-24]. Thus, it is the most widely used MHA quality rating tool in the scientific community. Furthermore, numerous international efforts promoting safe MHA use (eg, Mobile Health App Database, PsyberGuide or App Script, Reachout, Kinds Helpline, Health Navigator, and Vic Health) are based on the MARS.

The original version of the MARS is in English, but culture- and language-specific app ratings are needed globally. Spanish and Italian versions of the MARS have been developed [25,26]. A German MARS is necessary considering the growing and unregulated MHA market in Germany. Therefore, this study aimed to develop and validate a German version of the Mobile App Rating Scale (MARS-G) and to investigate the scalability of the overall MARS score with Mokken scale analysis—an approach that is closely related to item response theory.

Methods

Adaptation and Translation

The MARS was translated from English into German by two independent bilingual scientists (EMM and TP). After review and discussion of both forward translations, a pilot version of the MARS-G was created. This pilot version underwent blind back translation by two bilingual speakers with different backgrounds (a postdoctoral psychologist [AB] and a nonacademic individual [LMZ]). Thereafter, the back translation was compared with the original English version by the bilingual scientists (EMM and TP), and the penultimate version of the MARS-G was created. This version was evaluated for comprehensibility by three researchers and three nonacademics. After addressing their comments, the final version of the MARS-G was created and used in this study. The MARS-G can

be downloaded from the supplementary materials or obtained from the authors on request.

Search and Procedure

The MARS-G was validated within the framework of a study on the quality of apps targeting anxiety (E M Messner et al, unpublished data, 2020). Apps were identified using the following search terms: anxiety, fears, anxiety attack, anxious, anxiousness, anxiety disorder, fear, dread, fearful, panic, panic attacks, worry, and worries. Each search term was provided separately, as no truncation or use of logic operators (AND, OR, and NOT) was possible in the Google Play Store and iOS Store.

The inclusion process was divided into three steps (searching, screening, and determining eligibility). (1) Using the search terms mentioned above, the initial app pool was identified. (2) App details on the store sites were screened, and apps were downloaded and reviewed if they were developed for anxiety, were available in German or English, were downloadable through the official Google Play Store or iOS Store, and met no relevant exclusion criteria (app bundles [many applications only available as a group]). (3) All downloaded apps were assessed and excluded if they did not address anxiety, were not in German or English, were malfunctioning, or met relevant exclusion criteria (device incompatibility and development/test phase). We identified 3562 MHAs from the app stores. However, we excluded 810 duplicate apps, 2577 apps considered inappropriate on screening, and 71 apps considered ineligible. The remaining 104 apps were rated using the MARS and MARS-G by two independent trained raters. The raters tested all MHAs for 15 minutes. Quality was assessed immediately after the testing period in both languages. The assessment of the MARS-G is present in a review evaluating the quality of MHAs available for anxiety (E M Messner et al, unpublished data, 2020).

Rater Training

We followed the rating methodology in the original study by Stoyanov and colleagues [16]. We created a YouTube video with an introduction on MARS-G rating and an exercise on how to rate an app used as an exemplary health app (TrackYourTinnitus) [27]. This video can be requested from the corresponding author. Each rater was trained using this video, and five predefined apps were then rated to ensure that each rater was appropriately trained. If the individual rating score was different from our standard rating score by at least 2 points, the difference was discussed until agreement. All raters had at least a bachelor's degree in psychology to ensure a necessary minimum psychodiagnostic competence standard.

German Version of the Mobile Application Rating Scale

We added the following items in the app description section for the MARS-G: theoretical background (cognitive-behavioral, therapy, systemic therapy, etc), methods (eye movement desensitization and reprocessing, tracking, feedback, etc), category in the app store (lifestyle, medicine, etc), embedding into routine care (communication with therapist, etc), type of use (prevention, treatment, rehabilitation, etc), guidance (stand-alone, blended care, etc), certification (medical device law, etc), and data safety (log in, informed consent, etc). The four sections of the original MARS were expanded with an additional section focusing on the therapeutic gain associated with the app. The derived items were as follows: gain for the patient; gain for the therapist; risks and adverse effects; and ease of implementation in routine health care.

Analyses

Intraclass Correlation

The included MHAs were rated independently by two trained raters. The intraclass correlation coefficient (ICC) was calculated to assess the extent of agreement between the raters. An ICC of <0.50 indicated poor correlation, 0.51-0.75 indicated moderate correlation, 0.76-0.89 indicated good correlation, and >0.90 indicated excellent correlation [28]. According to the findings of previous studies, an ICC >0.75 was considered to indicate sufficient correlation [8,29,30].

Internal Consistency

Internal consistency of the MARS-G and its subscales was assessed as a measure of scale reliability, similar to the original MARS [16]. Omega was used instead of the widely adopted Cronbach alpha to assess reliability, as it provides a more unbiased estimation of reliability [31-33]. For estimations, the procedure introduced by Zhang and Yuan [34] was used to obtain robust coefficients and bootstrapped bias-corrected confidence intervals. Reliability of ω <0.50 indicated unacceptable internal consistency, 0.51-0.59 indicated poor consistency, 0.60-0.69 indicated questionable consistency, 0.70-0.79 indicated acceptable consistency, 0.80-0.89 indicated good consistency, and >0.90 indicated excellent consistency [35].

Validity

We assessed correlations between corresponding subscales of the MARS and MARS-G, as well as the overall correlation between the MARS total score and MARS-G total score. A *r* value >0.8 was a priori considered by the author group as an indicator of a strong and sufficient association between the MARS and MARS-G. Additionally, mean comparisons were performed between the dimensions of the MARS and MARS-G, using two-sided *t* tests. For all comparisons, a *P* value <.05 was considered significant.

Mokken Scale Analysis

Mokken scale analysis (MSA) is a scaling approach closely related to nonparametric item response theory [36]. The preconditions to use MSA are monotonicity and nonintersection. The key parameter in the MSA is Loevinger H. H_i is the scaling parameter for item i, and the overall scalability of all items clustering onto scale k is H_k. H_i indicates the strength of the relationship between a latent variable (app quality) and item i. A high scalability score indicates a high probability that an increase in item i is accompanied by an increase in the latent variable. A scale is considered weak if H is <0.4, moderate if H is ≥0.4 but <0.49, and strong if H is >0.5 [36]. This approach has been described in detail previously [36-39]. For both the MARS and MARS-G, the MSA was conducted to assess the

scalability of the mean scores. As recommended by van der Ark [36], the reliability of the scales was additionally assessed using the Molenaar-Sijtsma method (MS) [40,41], lambda-2 [42], and latent class reliability coefficient (LCRC) [43]. The MSA has been described previously [36].

Analysis Software

R software (R Foundation for Statistical Computing, Vienna, Austria) [44] was used for all analyses, except intraclass correlation. The MSA was conducted using the R package *mokken* [36,38]. Correlations and internal consistency were calculated using the *psych* (version 1.8.12) [45] and *coefficientalpha* packages (version 0.5) [34]. The *coefficientalpha* package includes the calculation of omega with missing and nonnormal data. The ICC was calculated using IBM SPSS 24 (IBM Corp, Armonk, New York) [46].

Results

Descriptive Data and Mean Comparisons

The ICCs for the MARS and MARS-G were high (ICC_{MARS}: 0.84, 95% CI 0.82-0.85; ICC_{MARS-G}: 0.83, 95% CI 0.82-0.85). The mean and standard deviation scores of the items in the MARS-G are presented in Table 1. The mean and standard deviation scores of the items in the MARS are reported elsewhere (E M Messner et al, unpublished data, 2020). The mean scores of the dimensions *engagement* (t_{206}=0.12; *P*=.91), *functionality* (t_{205}=0.39; *P*=.70), *esthetics* (t_{206}=−0.012; *P*=.99), and *information quality* (t_{204}=0.45; *P*=.66) and the overall rating (t_{206}=0.27; *P*=.80) were equivalent between the MARS and MARS-G.

Table 1. Summary of item and scale scores for the German version of the Mobile App Rating Scale.

Dimension	Score, mean (SD)
Engagement	**2.52 (0.70)**
Item 01	2.64 (0.93)
Item 02	2.79 (0.90)
Item 03	2.19 (1.00)
Item 04	1.86 (0.79)
Item 05	3.15 (0.72)
Functionality	**4.12 (0.69)**
Item 06	4.13 (0.82)
Item 07	4.24 (0.77)
Item 08	4.09 (0.74)
Item 09	4.03 (0.78)
Esthetics	**3.21 (0.94)**
Item 10	3.40 (0.93)
Item 11	3.20 (1.09)
Item 12	3.04 (0.99)
Information quality	**2.75 (0.60)**
Item 13	3.60 (0.76)
Item 14	2.63 (0.68)
Item 15	2.67 (0.76)
Item 16	2.61 (0.88)
Item 17	3.66 (0.68)
Item 18	1.87 (0.89)
Item 19	3.00 (N/A[a])
Overall mean	3.11 (0.58)

[a]This item on information quality could be rated for only 1 app, for the rest it was rated *not applicable*.

Internal Consistency

The internal consistency for the MARS dimension *engagement* was good (ω=0.84, 95% CI 0.77-0.88). The internal consistencies for *functionality* (ω=0.90, 95% CI 0.85-0.94) and *esthetics* (ω=0.91, 95% CI 0.92-0.96) were excellent. The internal consistency for *information quality* was acceptable (ω=0.74, 95% CI 0.14-0.99; α=.75, 95% CI 0.67-0.83). The internal consistency of the overall MARS score was good (ω=0.81, 95% CI 0.74-0.86).

The internal consistencies of the MARS-G dimensions were almost identical to those of the original MARS (engagement: ω=0.85, 95% CI 0.78-0.89; functionality: ω=0.91, 95% CI 0.87-0.94; esthetics: ω=0.93, 95% CI 0.90-0.95; information quality: ω=0.72, 95% CI 0.33-0.81). The internal consistency of the overall score was good (ω=0.82, 95% CI 0.76-0.86).

Validity

The correlation coefficients between corresponding dimensions of the MARS and MARS-G ranged from 0.93 to 0.98, and *P* values were adjusted for multiple testing according to the Holmes method [47] (Table 2). Correlations between the respective items are presented in Multimedia Appendix 1. There were no associations between user ratings and quality ratings (Table 1).

Table 2. Validity of the German version of the Mobile App Rating Scale (r and *P* value).

Dimension	$Engagement_{GER}$[a]	$Functionality_{GER}$	$Esthetics_{GER}$	$Information\ quality_{GER}$	Star rating
$Engagement_{ENG}$[b]	0.97 (<.001)	0.49 (<.001)	0.73 (<.001)	0.52 (.001)	−0.03 (.99)
$Functionality_{ENG}$	0.45 (<.001)	0.98 (<.001)	0.43 (<.001)	0.36 (.002)	0.06 (.99)
$Esthetics_{ENG}$	0.69 (<.001)	0.41 (<.001)	0.97 (<.001)	0.41 (.001)	0.12 (.99)
$Information\ quality_{ENG}$	0.55 (<.001)	0.34 (.004)	0.47 (<.001)	0.93 (.001)	0.25 (.19)
Star rating	−0.03 (>.99)	0.07 (>.99)	0.12 (>.99)	0.26 (.19)	—[c]

[a]German version.

[b]English version.

[c]Not applicable.

Mokken Scale Analysis

The MSA of the MARS revealed strong scalability (H=0.50; SE 0.062). There were no violations of monotonicity and nonintersection. The internal consistency of this scale was acceptable (MS=0.74; lambda 2=0.73; LCRC=0.72). The MSA of the MARS-G revealed good scalability (H=0.48; SE 0.060). The internal consistency of this scale was acceptable (MS=0.74; lambda 2=0.72; LCRC=0.74). The scalability results of the MARS and MARS-G are presented in Table 3.

Table 3. Summary of the Hk coefficient (overall scalability of all items in the scale) for the Mobile App Rating Scale (MARS) and the German version of the Mobile App Rating Scale (MARS-G).

Dimension	MARS	MARS-G
Engagement	0.59	0.57
Functionality	0.43	0.41
Esthetics	0.51	0.51
Information quality	0.45	0.41
Total scale	0.50	0.48

Discussion

Principal Findings

This study developed and evaluated the MARS-G for MHAs. The results showed that the MARS-G is a reliable and valid tool for experts to assess the quality of MHAs. The validity and reliability of the MARS-G were comparable to those of the original MARS. With regard to the reliability of the dimension *information quality*, the confidence interval of omega was overestimated owing to planned missingness. The planned missingness originated from the response option *not applicable*, which allows raters to skip an item if the app does not have any health information (eg, diary apps and brain games). There were no differences in reliability between the MARS-G and original MARS.

The MSA revealed that the use of the MARS-G total score is appropriate. Furthermore, there was good correspondence between the MARS-G and original MARS, indicating good validity. Our results are consistent with the findings of a study that introduced and tested an Italian version of the MARS [25].

The MARS-G has been presented in Multimedia Appendix 2 and can be obtained from the authors on request. It can be used freely for research and noncommercial MHA-evaluation projects. To reach satisfactory interrater reliability, completion of an online training exercise provided by the corresponding author is highly recommended. Furthermore, a training dataset of five apps can be obtained from the corresponding author on request. The MARS-G ratings should be revised until an appropriate level (ie, ICC >0.75) of interrater reliability is achieved.

To assist in MHA selection, standardized high-quality ratings of MHA are needed in German-speaking countries. Overall, a publicly available database presenting reliable, valid, and standardized expert ratings, like MARS-G ratings, could contribute to informed health care decisions on which app to

use for a specific disease or purpose. The mobile health app database [48] is one example of such a tool that assists users and health care providers in selecting appropriate apps for different health-related purposes.

Limitations

This study has several limitations. First, convergent validity was only evaluated by comparing the MARS and MARS-G. Comparisons with other app rating scales, such as ENLIGHT [15] and the American Psychological Association app evaluation model [12], are necessary in future studies. Second, the focus on anxiety apps limits generalization. Further studies are needed to confirm that these findings can be generalized to other mobile health domains. Such studies would require expert raters who are familiar with the specific domain. Finally, a confirmatory factor analysis of the MARS and MARS-G should be conducted in future studies with larger samples to ensure that the predefined subscales of the MARS and MARS-G can be confirmed.

Future Research

This translation study of the MARS led to the discovery of several research gaps. Future studies should focus on the improvement of app quality assessment and therefore the augmentation of safe MHA use on a broad scale. A challenge in this research is that the sequence in which apps are presented in the app store is incomprehensible and differs depending on which account is used for the search. In future studies, a web crawler could be used to search European app stores with keywords in order to build an unbiased database of available MHAs. Such a database already exists in China, and it contains all MHAs available in the United States, China, Japan, Brazil, and Russia [49].

Future studies should also shed light on the correlation between real-life user behavior and MARS or MARS-G ratings. As the MARS and MARS-G capture app quality, they could help predict the ability of users to download and use digital resources. Such research has already been conducted for ENLIGHT and real-life user engagement [50]. The efficacy of MHAs is strongly related to user adherence [50-52]; thus, high-quality apps might need to include adherence facilitation strategies to reach their potential.

Moreover, patient involvement should be taken into account. The user version of the MARS (uMARS) [53] should be translated and tested for reliability and validity as well, so that expert ratings of the MARS-G can be complemented with user ratings of the uMARS-G in German-speaking countries. In addition, there is a need for additional studies in the future to investigate the MARS-G and uMARS-G for apps related to specific health problems.

In conclusion, the MARS-G could be used by various stakeholders, such as public health authorities, patient organizations, researchers, health care providers (eg, physicians and psychotherapists), and interested third parties, to assess MHA quality. Furthermore, app developers could use the MARS-G as a tool to improve the quality of their apps.

Acknowledgments

The authors thank Linda Maria Zisch for her help in the translation process.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Item correlation matrix of MARS and MARS-G.
[DOCX File , 25 KB - mhealth_v8i3e14479_app1.docx]

Multimedia Appendix 2
Mobile Application Rating Scale-German.
[PDF File (Adobe PDF File), 563 KB - mhealth_v8i3e14479_app2.pdf]

References

1. UM London, eMarketer. Statista - The Statistics Portal. 2014 Aug. Smartphone user penetration as percentage of total population in Western Europe from 2011 to 2018 URL: https://www.statista.com/statistics/203722/smartphone-penetration-per-capita-in-western-europe-since-2000/ [accessed 2019-12-05]
2. ForwardAdGroup. Statista - Das Statistik-Portal. 2015. Wie viele Apps haben Sie auf Ihrem Smartphone installiert? URL: https://de.statista.com/statistik/daten/studie/162374/umfrage/durchschnittliche-anzahl-von-apps-auf-dem-handy-in-deutschland/ [accessed 2019-12-05]
3. Thranberend T, Knöppler K, Neisecke T. Gesundheits-Apps: Bedeutender Hebel für Patient Empowerment - Potenziale jedoch bislang kaum genutzt. Spotlight Gesundh 2016;2:1-8.
4. Deutschland.de. Deutschland.de. 2018. We speak German URL: https://www.deutschland.de/en/topic/culture/the-german-language-surprising-facts-and-figures [accessed 2019-04-24]

5. Contributors Wikipedia. Wikipedia, The Free Encyclopedia. 2019. List of languages by number of native speakers URL: https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers [accessed 2019-04-24]
6. Ellis E, Gogolin I, Clyne M. The Janus Face of Monolingualism: A Comparison of German and Australian Language Education Policies. Curr Issues Lang Plan 2010;11(4):439-460. [doi: 10.1080/14664208.2010.550544] [Medline: 26281194]
7. Heron KE, Smyth JM. Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
8. Terhorst Y, Rathner EM, Baumeister H, Sander L. «Hilfe aus dem App-Store?»: Eine systematische Übersichtsarbeit und Evaluation von Apps zur Anwendung bei Depressionen. Verhaltenstherapie 2018 May 8;28(2):101-112. [doi: 10.1159/000481692]
9. Ebert DD, Van Daele T, Nordgreen T, Karekla M, Compare A, Zarbo C, et al. Internet and mobile-based psychological interventions: Applications, efficacy and potential for improving mental health. Eur Psychol 2018 Jul;23(2):167-187. [doi: 10.1027/1016-9040/a000346]
10. Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014 Feb;5(3):229 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4814] [Medline: 24683442]
11. Albrecht UV. Kapitel 8. Gesundheits-Apps und Risiken. In: Albrecht UV, editor. Chancen und Risiken von Gesundheits-Apps (CHARISMHA). Hannover: Medizinische Hochschule Hannover; 2016:176-192.
12. American Psychiatric Association. American Psychiatric Association. 2017. App evaluation model URL: https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model [accessed 2019-12-05]
13. 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;4:363-371 [FREE Full text] [doi: 10.1007/s13142-014-0293-9] [Medline: 25584085]
14. Nouri R, Kalhori S, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. J Am Med Informatics Assoc 2018:1-10 [FREE Full text] [doi: 10.1093/jamia/ocy050]
15. Baumel A, Faber K, Mathur N, Kane JM, Muench F. Enlight: A Comprehensive Quality and Therapeutic Potential Evaluation Tool for Mobile and Web-Based eHealth Interventions. J Med Internet Res 2017 Mar 21;19(3):e82 [FREE Full text] [doi: 10.2196/jmir.7270] [Medline: 28325712]
16. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR mHealth uHealth 2015 Mar;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
17. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016 Mar 10;13:35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
18. Grainger R, Townsley H, White B, Langlotz T, Taylor WJ. Apps for People With Rheumatoid Arthritis to Monitor Their Disease Activity: A Review of Apps for Best Practice and Quality. JMIR Mhealth Uhealth 2017 Feb 21;5(2):e7 [FREE Full text] [doi: 10.2196/mhealth.6956] [Medline: 28223263]
19. Knitza J, Tascilar K, Messner EM, Meyer M, Vossen D, Pulla A, et al. German Mobile Apps in Rheumatology: Review and Analysis Using the Mobile Application Rating Scale (MARS). JMIR Mhealth Uhealth 2019 Aug 05;7(8):e14991 [FREE Full text] [doi: 10.2196/14991] [Medline: 31381501]
20. Machado GC, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: A systematic review. Best Practice & Research Clinical Rheumatology 2016 Dec;30(6):1098-1109. [doi: 10.1016/J.BERH.2017.04.002]
21. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and Evaluation of Mindfulness-Based iPhone Apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
22. Masterson Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and Analysis of Existing Mobile Phone Apps to Support Heart Failure Symptom Monitoring and Self-Care Management Using the Mobile Application Rating Scale (MARS). JMIR Mhealth Uhealth 2016 Jun 14;4(2):e74 [FREE Full text] [doi: 10.2196/mhealth.5882] [Medline: 27302310]
23. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the Quality of Mobile Apps for the Management of Pain: Systematic Search and Evaluation Using the Mobile App Rating Scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718 [FREE Full text] [doi: 10.2196/10718] [Medline: 30361196]
24. Thornton L, Quinn C, Birrell L, Guillaumier A, Shaw B, Forbes E, et al. Free smoking cessation mobile apps available in Australia: a quality review and content analysis. Aust N Z J Public Health 2017 Dec;41(6):625-630. [doi: 10.1111/1753-6405.12688] [Medline: 28749591]
25. Domnich A, Arata L, Amicizia D, Signori A, Patrick B, Stoyanov S, et al. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. BMC Med Inform Decis Mak 2016 Jul 7;16(83):1-10. [doi: 10.1186/s12911-016-0323-2]

26. Martin Payo R, Fernandez Álvarez MM, Blanco Díaz M, Cuesta Izquierdo M, Stoyanov SR, Llaneza Suárez E. Spanish adaptation and validation of the Mobile Application Rating Scale questionnaire. Int J Med Inform 2019 Sep;129:95-99. [doi: 10.1016/j.ijmedinf.2019.06.005]
27. Pryss R, Probst T, Schlee W, Schobel J, Langguth B, Neff P, et al. Prospective crowdsensing versus retrospective ratings of tinnitus variability and tinnitus–stress associations based on the TrackYourTinnitus mobile platform. Int J Data Sci Anal 2019:327-338. [doi: 10.1007/s41060-018-0111-4]
28. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Upper Saddle River, NJ: Pearson/Prentice Hall; 2009.
29. Lin J, Sander L, Paganini S, Schlicker S, Ebert D, Berking M, et al. Effectiveness and cost-effectiveness of a guided internet- and mobile-based depression intervention for individuals with chronic back pain: Protocol of a multi-centre randomised controlled trial. BMJ Open 2017 Dec 28;7:e015226. [doi: 10.1136/bmjopen-2016-015226]
30. Sander L, Paganini S, Lin J, Schlicker S, Ebert DD, Buntrock C, et al. Effectiveness and cost-effectiveness of a guided Internet- and mobile-based intervention for the indicated prevention of major depression in patients with chronic back pain—study protocol of the PROD-BP multicenter pragmatic RCT. BMC Psychiatry 2017 Jan 21;17(36). [doi: 10.1186/s12888-017-1193-6]
31. Dunn TJ, Baguley T, Brunsden V. From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. Br J Psychol 2014 Aug;105(3):399-412. [doi: 10.1111/bjop.12046] [Medline: 24844115]
32. Revelle W, Zinbarg RE. Coefficients Alpha, Beta, Omega, and the GLB: Comments on Sijtsma. Psychometrika 2009;74(1):145-154. [doi: 10.1007/s11336-008-9102-z]
33. McNeish D. Thanks Coefficient Alpha, We'll Take It From Here. Psychol Methods 2018 Sep;23(3):412-433. [doi: 10.1037/met0000144]
34. Zhang Z, Yuan KH. Robust Coefficients Alpha and Omega and Confidence Intervals With Outlying Observations and Missing Data: Methods and Software. Educ Psychol Meas 2016 Jun;76(3):387-411 [FREE Full text] [doi: 10.1177/0013164415594658] [Medline: 29795870]
35. George D, Mallery P. SPSS For Windows Step By Step: A Simple Guide And Reference, 11.0 Update. Boston: Allyn & Bacon; 2003.
36. van der Ark LA. Mokken Scale Analysis in R. J Stat Softw 2007 Nov 8;20(11):1-19. [doi: 10.1007/s11336-007-9034-z]
37. Mokken RJ. A theory and procedure of scale analysis: With applications in political research. New York: De Gruyter Mouton; 1971.
38. van der Ark LA. New Developments in Mokken Scale Analysis in R. J Stat Softw 2012;48(5):1-27. [doi: 10.18637/jss.v048.i05]
39. Sijtsma K, van der Ark LA. A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. Br J Math Stat Psychol 2017;70(1):137-158. [doi: 10.1111/bmsp.12078]
40. Molenaar I, Sijtsma K. Mokken's Approach to Reliability Estimation Extended to Multicategory Items. Kwant Methoden 1988;9(28):115-126.
41. Sijtsma K, Molenaar IW. Reliability of test scores in nonparametric item response theory. Psychometrika 1987 Mar;52(1):79-97. [doi: 10.1007/bf02293957]
42. Guttman L. A basis for analyzing test-retest reliability. Psychometrika 1945 Dec;10(4):255-282. [doi: 10.1007/bf02288892]
43. van der Ark LA, van der Palm DW, Sijtsma K. A Latent Class Approach to Estimating Test-Score Reliability. Appl Psychol Meas 2011 Mar 09;35(5):380-392. [doi: 10.1177/0146621610392911]
44. R Core Team. R: A Language and Environment for Statistical Computing. R Found Stat Comput Vienna, Austria 2017 [FREE Full text]
45. Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research [Computer Software]. 2017. URL: https://personality-project.org/r/psych [accessed 2019-11-22]
46. IBM Corporation. IBM SPSS Advanced Statistics 24 [Software]. 2016. URL: http://www-01.ibm.com/support/docview.wss?uid=swg27047033#ja%5Cnftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/24.0/ja/client/Manuals/IBM_SPSS_Advanced_ [accessed 2020-01-16]
47. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat 1979;6:65-70 [FREE Full text]
48. MHAD Core Team. Mobile Health App Database. 2019. URL: http://www.mhad.science/ [accessed 2019-09-11]
49. Xu W, Liu Y. mHealthApps: A Repository and Database of Mobile Health Apps. JMIR mHealth uHealth 2015 Mar 18;3(1):e28 [FREE Full text] [doi: 10.2196/mhealth.4026] [Medline: 25786060]
50. Baumel A, Yom-Tov E. Predicting user adherence to behavioral eHealth interventions in the real world: examining which aspects of intervention design matter most. Transl Behav Med 2018 Sep 08;8(5):793-798. [doi: 10.1093/tbm/ibx037] [Medline: 29471424]
51. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression: Systematic review. J Med Internet Res 2009 Apr;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
52. Van Ballegooijen W, Cuijpers P, Van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to internet-based and face-to-face cognitive behavioural therapy for depression: A meta-analysis. PLoS ONE 2014 Jul;9(7):e100674 [FREE Full text] [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]

53. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS). JMIR mHealth uHealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]

Abbreviations

ENLIGHT: Evaluation Tool for Mobile and Web-Based eHealth Interventions
ICC: intraclass correlation coefficient
LCRC: latent class reliability coefficient
MARS: Mobile App Rating Scale
MARS-G: German version of the Mobile App Rating Scale
MHA: mobile health app
MS: Molenaar-Sijtsma method
MSA: Mokken scale analysis
uMARS: user version of the Mobile App Rating Scale

Edited by G Eysenbach; submitted 24.04.19; peer-reviewed by C Aljoscha, M Bardus, E de Krijger, R Bipeta; comments to author 05.06.19; revised version received 29.07.19; accepted 24.09.19; published 27.03.20.

Please cite as:
Messner EM, Terhorst Y, Barke A, Baumeister H, Stoyanov S, Hides L, Kavanagh D, Pryss R, Sander L, Probst T
The German Version of the Mobile App Rating Scale (MARS-G): Development and Validation Study
JMIR Mhealth Uhealth 2020;8(3):e14479
URL: http://mhealth.jmir.org/2020/3/e14479/
doi:10.2196/14479
PMID:32217504

Review

Long-Term Weight Management Using Wearable Technology in Overweight and Obese Adults: Systematic Review

Emily Fawcett[1]; Michelle Helena Van Velthoven[1], BSc, MSc, PhD; Edward Meinert[1,2], MA, MSc, MBA, MPA, PhD, CEng FBCS, EUR ING

[1]Department of Paediatrics, University of Oxford, Oxford, United Kingdom

[2]Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

Corresponding Author:
Edward Meinert, MA, MSc, MBA, MPA, PhD, CEng FBCS, EUR ING
Department of Paediatrics
University of Oxford
John Radcliffe Hospital
Children's Hospital
Oxford, OX3 9DU
United Kingdom
Phone: 44 7824446808
Email: e.meinert14@imperial.ac.uk

Abstract

Background: Although there are many wearable devices available to help people lose weight and decrease the rising prevalence of obesity, the effectiveness of these devices in long-term weight management has not been established.

Objective: This study aimed to systematically review the literature on using wearable technology for long-term weight loss in overweight and obese adults.

Methods: We searched the following databases: Medical Literature Analysis and Retrieval System Online, EMBASE, Compendex, ScienceDirect, Cochrane Central, and Scopus. The inclusion criteria were studies that took measurements for a period of ≥1 year (long-term) and had adult participants with a BMI >24. A total of 2 reviewers screened titles and abstracts and assessed the selected full-text papers for eligibility. The risk of bias assessment was performed using the following tools appropriate for different study types: the Cochrane risk of bias tool, Risk Of Bias In Nonrandomized Studies-of Interventions, A MeaSurement Tool to Assess systematic Reviews, and 6 questions to trigger critical thinking. The results of the studies have been provided in a narrative summary.

Results: We included five intervention studies: four randomized controlled trials and one nonrandomized study. In addition, we used insights from six systematic reviews, four commentary papers, and a dissertation. The interventions delivered by wearable devices did not show a benefit over comparator interventions, but overweight and obese participants still lost weight over time. The included intervention studies were likely to suffer from bias. Significant variances in objectives, methods, and results of included studies prevented meta-analysis.

Conclusions: This review showed some evidence that wearable devices can improve long-term physical activity and weight loss outcomes, but there was not enough evidence to show a benefit over the comparator methods. A major issue is the challenge of separating the effect of decreasing use of wearable devices over time from the effect of the wearable devices on the outcomes. Consistency in study methods is needed in future long-term studies on the use of wearable devices for weight loss.

Trial Registration: PROSPERO CRD42018096932; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=96932

(JMIR Mhealth Uhealth 2020;8(3):e13461) doi:10.2196/13461

KEYWORDS

telemedicine; mHealth; eHealth; mobile health; obesity; wearable electronic devices; wearable technology; wearable device; digital technology; weight loss; overweight; fitness trackers

Introduction

Obesity is a rising concern worldwide [1]. By 2030, obesity prevalence in the United States is predicted to be 50% to 51% in men and 45% to 52% in women, and it is estimated that in the United Kingdom, 41% to 48% of men and 35% to 43% of women will be obese [2]. Obesity is well known to be a risk factor for different medical conditions, leading to increased morbidity and mortality [2,3]. Various interacting factors influence the prevalence of obesity, including people's upbringing, lifestyle, environment, and genetics [4]. Over the past decades, numerous strategies for losing weight have been developed that mainly focus on reducing calorie intake and increasing energy expenditure [1]. It is important to tackle obesity early on, as the ability of a person to increase his or her activity levels decreases as his or her weight increases (particularly BMI >40) [4].

The rapid development of technology has led to a growing market of wearable devices claiming to help people lose weight. Over 100 million wearable devices were sold in 2016, and sales were expected to continue to rise over the next years [5]. Wearable technology refers to any electronic device that is worn on the body, commonly being fitness trackers containing some form of an activity monitor.

In combination with an effective weight management intervention based on a behavior change model, wearable technologies can help people lose weight through various means, eg, by promoting physical exercise, by monitoring food consumption, or by encouraging interuser communication and support [6]. Research on the effectiveness of interventions delivered by wearable devices suggests that these interventions can help lose weight [7]. However, long-term weight loss (>1 year) is often unsuccessful [8]. Wearable devices have only demonstrated a statistically significant weight loss lasting for a few weeks, which greatly reduces the potential usefulness of these devices [7]. Digital wearables could be a novelty that wears off over time, rather than being part of a sustained lifestyle change [9].

Previous research on weight loss interventions without wearable technology has shown that over a 5-year period, only 20% of individuals maintained a weight loss of more than 5 kg (after an initial loss of around 10 kg) [10,11]. Therefore, this review focused on studies that can aid long-term weight loss. Evidence on the long-term effects of wearables to manage or prevent obesity could be relevant for people seeking to reach a healthy weight and for their medical practitioners [12].

This study systematically reviewed the use of wearable devices for long-term weight loss in overweight and obese adults. This review had four objectives: (1) to investigate the effects of using wearable devices on physical activity and weight outcomes, (2) to examine the duration of wearable technology use, (3) to assess the accuracy of wearable technology vs self-reporting, and (4) to explore the use of wearable technology by people with specific medical conditions.

Methods

Protocol

A protocol was registered with the International Prospective Register for Systematic Reviews (CRD42018096932), with the review structure following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Multimedia Appendix 1). We narrowed down the review question of the protocol, focusing on long-term weight management.

Eligibility Criteria

Textbox 1 summarizes the inclusion and exclusion criteria for the participant, intervention, comparators, outcomes, and study types of this systematic review.

Textbox 1. Inclusion and exclusion criteria for study selection.

- Population: We included studies with obese or overweight adult participants. *Overweight* was defined as having a BMI ranging between 25 kg/m2 and 29.99 kg/m2 or as defined by the study. *Obese* was defined as having a BMI of 30 kg/m2 or more. Participants in hospital settings were excluded as these studies were unlikely to focus on long-term effects and as they do not represent the real-world use of wearable devices.
- Intervention: Interventions included digital wearable technologies used for monitoring or managing weight. Studies that only included a mobile phone app were excluded.
- Comparators: Comparators included traditional behavioral weight loss approaches, usual care, or another intervention. Studies that did not have a comparator were also included if they met the other inclusion criteria.
- Outcomes: The primary outcome was change in physical activity and weight after using digital wearable technology for at least a year. Secondary outcomes were the duration of wearable technology use, the accuracy of wearable technology vs self-reporting, and the use of wearable technology by people with specific medical conditions.
- Study types: All types of studies were included. Owing to the rapid advances in technology, studies from only the past 10 years were used—from 2008 onward.

Information Sources and Search

We searched the following databases: Medical Literature Analysis and Retrieval System Online, Compendex, ScienceDirect, Cochrane Central, Scopus, and EMBASE through Ovid. Multimedia Appendix 2 outlines the search terms. Data published before 2008 were not included as these data are not reflective of the rapid change in the use of mobile phones and wearables. Keywords related to participant, intervention, comparators, and outcome items were used to search for relevant papers. A librarian was consulted for advice on the searches. The search was adjusted and modified for each database.

Study Selection

The references found were imported into EndNote X9 (Clarivate, Pennsylvania), and duplicates were removed. Overall, 2 reviewers conducted title and abstract screening. Any differences in the chosen studies were discussed until a consensus was reached. The full texts of potentially eligible studies were retrieved and analyzed for eligibility by 2 reviewers.

Data Collection Process and Items

A standardized data extraction sheet was used to extract data. The extracted data included the title, the research question, the data sources, how the data were analyzed, the main findings, and the conclusions.

Quality Appraisal of Individual Studies

All included studies underwent a methodological quality appraisal. Relevant appraisal tools were used for different study designs. The Cochrane risk of bias tool was used for randomized controlled trials (RCTs), Risk Of Bias In Nonrandomized Studies-of Interventions was used for nonrandomized studies of interventions, A MeaSurement Tool to Assess systematic Reviews was used for systematic reviews, and *6 questions to trigger critical thinking* were used for qualitative papers [13-16].

Synthesis of Results

We have provided a narrative overview and tabular summary of the findings. A meta-analysis of the studies could not be conducted because of the heterogeneity in their interventions, participants, and outcomes.

Results

Study Selection and Characteristics

We found 1116 references, and after removing duplicates and adding six references identified through searching reference lists of included studies, 684 titles and abstracts were screened (Figure 1). Furthermore, 44 full texts were assessed for inclusion, of which 28 were excluded (Multimedia Appendix 3). We included five intervention studies: four RCTs and one nonrandomized study of an intervention as shown in Table 1. In addition, we used insights from six systematic reviews, four commentary papers, and a dissertation for additional insights (Table 2).

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram showing the selection of included studies.

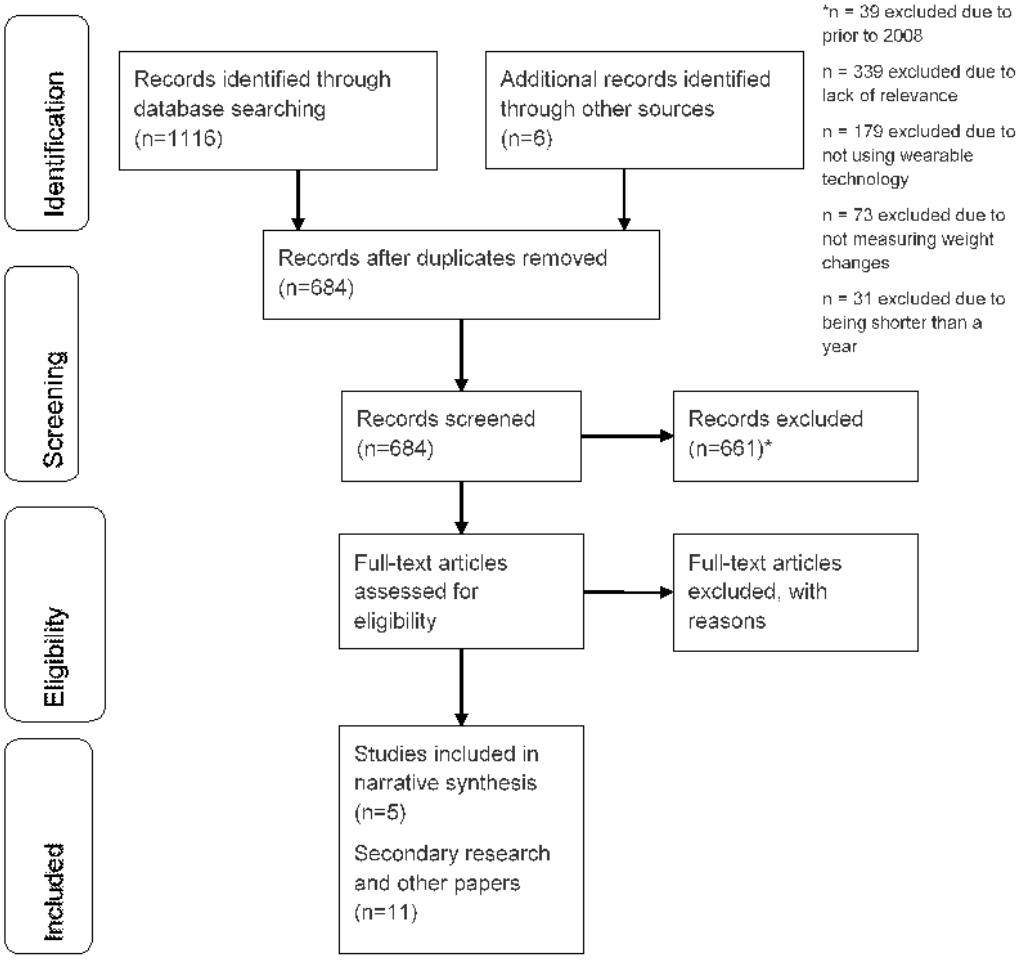


Table 1. Included intervention studies.

Study	Author (year)	Title	Type of paper
1	Fazzino et al (2017) [17]	Change in physical activity during a weight management intervention for breast cancer survivors: Association with weight outcomes	Randomized controlled trial
2	Chiang et al (2017) [18]	Potential impact of wearable technology as part of a multidisciplinary treatment strategy for weight regain following RYGB[a]	Randomized controlled trial
3	Jakicic et al (2016) [19]	Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: The IDEA randomized clinical trial	Randomized controlled trial
4	Jakicic et al (2012) [20]	Effect of a stepped-care intervention approach on weight loss in adults: A randomized clinical trial	Randomized controlled trial
5	Sepah et al (2015) [21]	Long-term outcomes of a Web-based diabetes prevention program: 2-year results of a single-arm longitudinal study	Nonrandomized study of an intervention

[a]RYGB: Roux-en-Y gastric bypass.

Table 2. Secondary research and other papers.

Study	Author (year)	Title	Type of paper
6	Podina and Fodor (2018) [16]	Critical review and meta-analysis of multicomponent behavioural e-health interventions for weight loss	Systematic review
7	Lyzwinski (2014) [22]	A systematic review and meta-analysis of mobile devices and weight loss with an intervention content analysis	Systematic review
8	Allen et al (2014) [23]	Technology-assisted weight management interventions: Systematic review of clinical trials	Systematic review
9	Kelders et al (2012) [24]	Persuasive system design does matter: A systematic review of adherence to Web-based interventions	Systematic review
10	Goode et al (2017) [25]	The impact of interventions that integrate accelerometers on physical activity and weight loss: A systematic review	Systematic review
11	Coons et al (2012) [26]	Technology interventions to curb obesity: A systematic review of the current literature	Systematic review
12	Kulick (2017) [27]	Wearable technology and long-term weight loss	Comment
13	Jakicic and Belle (2017) [28]	Wearable technology and long-term weight loss-Reply	Comment
14	Hekler et al (2016)	Advancing models and theories for digital behavior change interventions [6]	Comment
15	Dyer (2016) [29]	Wearable fitness device does not help maintain weight loss, study finds: Fitness device doesn't maintain weight loss	Comment
16	Assar (2018) [30]	Evidence-based psychotherapeutic interventions and mHealth for weight management in overweight: A biopsychosocial framework	Dissertation

Risk of Bias Within Studies

There was a large variation in bias in all the included papers (Multimedia Appendix 4). In most of the RCTs, there was a high risk of bias because of the lack of blinding of participants, as a blinded version of a wearable device intervention is not possible. All the included RCTs used random sequence generation to generate groups, although information about allocation concealment was missing in half of them.

Synthesis of Results

Table 3 outlines the studies that contributed findings to the fulfillment of the objectives of this review: physical activity and weight outcomes, duration of wearable technology use, accuracy of wearable technology over self-reporting, and use by people with specific medical conditions.

Table 3. Included intervention studies and findings.

Study	Author (year)	Description	Physical activity and weight outcomes	Long-term use	Accuracy compared with self-reporting	Population with a specific medical condition
1	Fazzino et al (2017) [17]	An RCT[a] assessing the effects of mobile health weight management on physical activity, weight loss, and weight maintenance	Yes	Yes	Yes	Breast cancer prevention
2	Chiang et al (2017) [18]	An RCT on the weight loss after repaired RYGB[b] surgery, with and without wearable devices	Yes	No	No	After repair of failed Roux-en-Y gastric bypass
3	Jakicic et al (2016) [19]	An RCT comparing outcomes of technology-enhanced interventions with standard behavioral interventions	Yes	Yes	Yes	No
4	Jakicic et al (2012) [20]	An RCT comparing a standard and stepped-care intervention in weight loss	No	Yes	No	No
5	Sepah et al (2015) [21]	A diabetes prevention study measuring the outcomes of weight and hemoglobin A_{1c}	No	Yes	No	Prediabetes

[a]RCT: randomized controlled trial.

[b]RYGB: Roux-en-Y gastric bypass.

Association Between Weight Outcomes and Change in Physical Activity

The three studies reporting on weight loss and physical activity outcomes concluded that using wearable devices had a benefit on these outcomes, but not compared with the comparator groups. Study 1 showed a significant rise in moderate-to-vigorous physical activity over 18 months and divided participants into high or low original weight loss and high or low weight regain groups. At 6 months, the high weight loss groups had significantly higher level of moderate-to-vigorous physical activity than the low weight loss group. However, at 12 and 18 months, the high loss and high regain groups' level of moderate-to-vigorous physical activity fell, leaving the high loss and low regain group with a significantly higher level of moderate-to-vigorous physical activity than all other groups.

Study 2 compared 27 individuals using wearable devices with 260 individuals who were not using a wearable device. A total of 8000 steps per day was recommended for the intervention group, but it is not noted whether the comparator group were given similar recommendations. A significant benefit was only found at 2 years (P=.03).

In both the standard and wearable device groups of study 3, there was an increase in the duration of moderate-to-vigorous physical activity sessions ≥10 min over a 12-month period, but there was no statistically significant difference between the two groups. The percentage of weight lost did differ, with significantly greater weight loss in the comparator group compared with the wearable devices group from 12 months onward. Study 3 did not find an association between physical activity and weight loss within its groups.

Maintenance of Wearable Technology Use

Retention was fairly high in four studies, but study 2 did not provide data (Table 4). Study 4 was the only study mentioning to offer a monetary incentive for assessments (US $10-$25). Study 4 compared a standard intervention with a stepped-care intervention, where the intensity of support (such as telephone intervention and additional individual sessions) increased if certain goals were not met.

Table 4. Retention rate across included intervention studies.

Study	Author (year)	Description	Retention at 6 months, %	Retention at 18 months, %	Retention at 24 months, %	Notes
1	Fazzino et al (2017) [17]	An RCT[a] assessing the effects of mobile health weight management on physical activity, weight loss, and weight maintenance	N/A[b]	68	N/A	80% maintained intervention use at 18 months but without valid accelerometer data.
2	Chiang et al (2017) [18]	An RCT on the weight loss after repaired RYGB[c] surgery, with and without wearable devices	N/A	N/A	N/A	N/A
3	Jakicic et al (2016 [19]	An RCT comparing outcomes of technology-enhanced interventions with standard behavioral interventions	N/A	N/A	75	N/A
4	Jakicic et al (2012) [20]	An RCT comparing a standard and stepped-care intervention in weight loss	N/A	72	N/A	N/A
5	Sepah et al (2015) [21]	A diabetes prevention study measuring the outcomes of weight and hemoglobin A_{1c}	79.1	N/A	70.1	N/A

[a]RCT: randomized controlled trial.

[b]N/A: not applicable.

[c]RYGB: Roux-en-Y gastric bypass.

Accuracy of Wearable Technology Versus Self-Reporting

Accuracy was reported by two studies. Study 1 found that self-reported moderate-to-vigorous physical activity was significantly higher than that recorded by the wearable device. The groups with poorer outcomes (low loss or high regain) had larger discrepancies between the two methods. The self-reported and accelerometer-derived moderate-to-vigorous physical activities were most similar to the high loss and low regain group. the moderate-to-vigorous physical activity data from self-reporting and the accelerometer were, however, collected on different weeks. However, the high loss and low regain group still overestimated moderate-to-vigorous physical activity, and the overestimation did not reduce over time. Study 1 suggested "social desirability to report physical activity adherence," with participants inflating self-reported moderate-to-vigorous physical activity.

Use of Wearable Technology by People With Specific Medical Conditions

Overall, three out of the five studies focused on populations with a specific medical condition (Table 4). This included a history of breast cancer (study 1), repair of failed Roux-en-Y gastric bypass (study 2), or prediabetes (study 5). Study 1 analyzed those who attended the visits but had invalid or missing data and found no significant difference in *cancer treatment–related variables*. Study 5 measured blood glucose (hemoglobin A_{1c}) levels and showed a significant beneficial reduction over 24 months.

Discussion

Principal Findings

This review showed some evidence that wearable devices can improve long-term physical activity and weight loss outcomes, but there was not enough evidence to show a benefit over the comparator methods. The comparator interventions differed among studies, which adds to the difficultly in determining the impact on outcomes. Although the term *standard* was used, there was no standardization in the comparators' intervention, with different levels of support and procedures.

Overall physical activity levels increased from baseline, but there was no difference between wearable and comparator interventions. Study 1 found that those who sustained higher physical activity levels were more likely to maintain weight loss. Retention was fairly high in the included intervention studies. The mechanism through which wearable devices have an effect compared with other methods was not known as diet and physical activity were not different. The accuracy of wearable devices varied, which could be explained by the different features and technology of wearables. A total of three included studies focused on populations with a specific medical condition. The difference in populations added a challenge to comparing the studies as the results of a study on the weight management of patients with one medical condition may not apply to patients with another medical condition or the general population.

Limitations

There were only five studies with a relatively small sample size assessing the long-term use of wearable devices. It was not possible to undertake a meta-analysis because of the heterogeneity among participants, wearables, methods, and

outcomes. The included studies were likely to suffer from bias. Wearable device interventions cannot be *blinded* to the user. Only outcome assessors could have been blinded, which most studies did not attempt to do. The use of wearables in these studies may not be applicable to real-world scenarios as the companies selling these wearable devices do not offer the support that was offered by researchers in the studies. A limitation of this review is that we only conducted a basic search limited to a few keywords and phrases. In addition, databases such as Institute of Electrical and Electronics Engineers Xplore and Cumulative Index of Nursing and Allied Health Literature for clinical and behavioral science research were not searched.

Comparison With Prior Work

It is important to retain participants in studies to separate the effect of the study design and intervention [31]. Study 5 compared a standard intervention with a stepped-care intervention where the intensity of support (such as telephone intervention and additional individual sessions) increased if certain goals were not met. Interventions of this kind have been shown to reduce attrition [32]. Other strategies for improving long-term data collection are offering incentives, reducing barriers by offering *alternative data collection modes*, and reminder calls [32]. Improving adoption and retention through methods such as monetary incentives could be counterproductive as this is not possible in real-life settings.

Consciously or subconsciously, self-reported physical activity levels are often overestimated [33]. Wearable devices are more accurate at estimating physical activity levels than self-reporting, though a truly objective method is currently not available for everyday purposes. Accelerometers, which were used in the wearable devices in the included studies, can lead to different estimates, even when using the same device [34]. The accuracy of heart rate monitors has been reported to be higher but still insufficient [35].

Wearable devices have shown benefit in managing medical conditions, eg, diabetes [36]. Studies with populations having medical conditions or risk factors could suffer from higher dropout because of the higher risk of a medical event. However, having a specific condition or medical event could be a stronger motivation than having a *vaguer* risk factor such as being overweight or obese [37].

Recommendations for Future Work

Different aspects of weight loss maintenance and wearable devices have been studied, but large areas are still unknown. These include the mechanisms through which using wearable devices can lead to weight loss and studies into the usefulness of wearable devices for long-term weight management.

Those who managed to sustain raised physical activity levels had a weight maintenance benefit. Not all groups managed to sustain increased activity levels, and it would be valuable to understand why. Individuals who sustained exercise could have been more likely to commit to other lifestyle changes around weight management.

Investigating the reasons for dropout could help to understand to what extent this is caused by study design and/or flaws in wearable devices. Discovering how wearable devices are being used, and whether their use is improved through outside support, would give valuable information for designing more effective wearable devices. It could also help health care practitioners to advice and support people who are trying to lose weight and are interested in using wearable devices.

Conclusions

We found a small number of long-term studies showing some evidence that wearable devices can improve long-term physical activity and weight loss outcomes, but there was not enough evidence to show a benefit over the comparator methods. A major issue is the challenge to separate the effect of the decreasing use of wearable devices over time from the effect of wearable devices on the outcomes. Consistency in study methods is needed in future long-term studies on the use of wearable devices for weight loss.

Acknowledgments

The authors would like to thank Karine Barker at the Radcliffe Science Library. Research issues identified and prioritized by the members of the public in a workshop at the European Scientific Institute in July 2017 were used to guide the focus of this study. The authors declare that the manuscript is an honest, accurate, and transparent account of the study being reported, that no important aspects of the study have been omitted, and that any discrepancies from the study as planned (and, if relevant, registered) have been explained. This study was a systematic review of publicly available literature. This work was supported by the Sir David Cooksey Fellowship in Healthcare Translation and the Final Honour School of Medical Sciences, Cell and Systems Biology and Neuroscience at the University of Oxford.

Authors' Contributions

EM conceived the study objectives and oversaw the original study protocol. EF reviewed the initial study protocol, made amendments as per this manuscript's methods, executed the review independently (with peer review on study inclusion), and drafted the final manuscript on her own. EM gave feedback to EF, and EF incorporated all feedback. DB also provided feedback on iterations. MV rewrote the paper and made major revisions based on peer-review feedback. All authors approved the final manuscript. EM is the guarantor.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Identify the report as a systematic review, meta-analysis, or both.
[DOCX File , 19 KB - mhealth_v8i3e13461_app1.docx]

Multimedia Appendix 2
Results found in each database and the search strings used.
[DOCX File , 13 KB - mhealth_v8i3e13461_app2.docx]

Multimedia Appendix 3
Excluded studies.
[DOCX File , 18 KB - mhealth_v8i3e13461_app3.docx]

Multimedia Appendix 4
Quality appraisal of randomized controlled trials.
[DOCX File , 41 KB - mhealth_v8i3e13461_app4.docx]

References

1. Caballero B. The global epidemic of obesity: an overview. Epidemiol Rev 2007;29:1-5. [doi: 10.1093/epirev/mxm012] [Medline: 17569676]
2. Wang YC, McPherson K, Marsh T, Gortmaker SL, Brown M. Health and economic burden of the projected obesity trends in the USA and the UK. Lancet 2011 Aug 27;378(9793):815-825. [doi: 10.1016/S0140-6736(11)60814-3] [Medline: 21872750]
3. Pi-Sunyer X. The medical risks of obesity. Postgrad Med 2009 Nov;121(6):21-33 [FREE Full text] [doi: 10.3810/pgm.2009.11.2074] [Medline: 19940414]
4. Forhan M, Gill SV. Obesity, functional mobility and quality of life. Best Pract Res Clin Endocrinol Metab 2013 Apr;27(2):129-137. [doi: 10.1016/j.beem.2013.01.003] [Medline: 23731875]
5. Wade J. Smart Insights. 2017 Nov 15. Wearable Technology Statistics and Trends 2018 URL: https://www.smartinsights.com/digital-marketing-strategy/wearables-statistics-2017/ [accessed 2019-08-10]
6. Hekler EB, Michie S, Pavel M, Rivera DE, Collins LM, Jimison HB, et al. Advancing models and theories for digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):825-832 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.013] [Medline: 27745682]
7. Cheatham SW, Stull KR, Fantigrassi M, Motel I. The efficacy of wearable activity tracking technology as part of a weight loss program: a systematic review. J Sports Med Phys Fitness 2018 Apr;58(4):534-548. [doi: 10.23736/S0022-4707.17.07437-0] [Medline: 28488834]
8. Anderson JW, Konz EC, Frederich RC, Wood CL. Long-term weight-loss maintenance: a meta-analysis of US studies. Am J Clin Nutr 2001 Nov;74(5):579-584. [doi: 10.1093/ajcn/74.5.579] [Medline: 11684524]
9. McGuire M, Wing R, Hill J. The prevalence of weight loss maintenance among American adults. Int J Obes Relat Metab Disord 1999 Dec;23(12):1314-1319. [doi: 10.1038/sj.ijo.0801075] [Medline: 10643690]
10. Wing RR, Hill JO. Successful weight loss maintenance. Annu Rev Nutr 2001;21:323-341. [doi: 10.1146/annurev.nutr.21.1.323] [Medline: 11375440]
11. Relias Media. 2016 Nov 1. Does Technology Accelerate Weight Loss? URL: https://www.reliasmedia.com/articles/138992-does-technology-accelerate-weight-loss [accessed 2020-02-05]
12. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Br Med J 2011 Oct 18;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
13. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. Br Med J 2016 Oct 12;355:i4919 [FREE Full text] [doi: 10.1136/bmj.i4919] [Medline: 27733354]
14. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. Br Med J 2017 Sep 21;358:j4008 [FREE Full text] [doi: 10.1136/bmj.j4008] [Medline: 28935701]
15. Aveyard H, Sharp P, Woolliams M. A Beginner's Guide to Critical Thinking and Writing in Health and Social Care. Maidenhead, Berkshire: Open University Press; 2015.
16. Podina IR, Fodor LA. Critical review and meta-analysis of multicomponent behavioral e-health interventions for weight loss. Health Psychol 2018 Jun;37(6):501-515. [doi: 10.1037/hea0000623] [Medline: 29733617]
17. Fazzino TL, Fabian C, Befort CA. Change in physical activity during a weight management intervention for breast cancer survivors: association with weight outcomes. Obesity (Silver Spring) 2017 Nov;25(Suppl 2):S109-S115 [FREE Full text] [doi: 10.1002/oby.22007] [Medline: 29086523]

18. Chiang AL, Jirapinyo P, Thompson CC. Potential impact of wearable technology as part of a multidisciplinary treatment strategy for weight regain following roux-en-y gastric bypass. Gastroenterology 2017;152(5):S1016-S1017. [doi: 10.1016/s0016-5085(17)33445-5]
19. 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. J Am Med Assoc 2016 Sep 20;316(11):1161-1171 [FREE Full text] [doi: 10.1001/jama.2016.12858] [Medline: 27654602]
20. Jakicic JM, Tate DF, Lang W, Davis KK, Polzien K, Rickman AD, et al. Effect of a stepped-care intervention approach on weight loss in adults: a randomized clinical trial. J Am Med Assoc 2012 Jun 27;307(24):2617-2626 [FREE Full text] [doi: 10.1001/jama.2012.6866] [Medline: 22735431]
21. Sepah SC, Jiang L, Peters AL. Long-term outcomes of a web-based diabetes prevention program: 2-year results of a single-arm longitudinal study. J Med Internet Res 2015 Apr 10;17(4):e92 [FREE Full text] [doi: 10.2196/jmir.4052] [Medline: 25863515]
22. Lyzwinski LN. A systematic review and meta-analysis of mobile devices and weight loss with an intervention content analysis. J Pers Med 2014 Jun 30;4(3):311-385 [FREE Full text] [doi: 10.3390/jpm4030311] [Medline: 25563356]
23. Allen JK, Stephens J, Patel A. Technology-assisted weight management interventions: systematic review of clinical trials. Telemed J E Health 2014 Dec;20(12):1103-1120 [FREE Full text] [doi: 10.1089/tmj.2014.0030] [Medline: 25409001]
24. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
25. Goode AP, Hall KS, Batch BC, Huffman KM, Hastings SN, Allen KD, et al. The impact of interventions that integrate accelerometers on physical activity and weight loss: a systematic review. Ann Behav Med 2017 Feb;51(1):79-93 [FREE Full text] [doi: 10.1007/s12160-016-9829-1] [Medline: 27565168]
26. Coons MJ, Demott A, Buscemi J, Duncan JM, Pellegrini CA, Steglitz J, et al. Technology interventions to curb obesity: a systematic review of the current literature. Curr Cardiovasc Risk Rep 2012 Apr;6(2):120-134 [FREE Full text] [doi: 10.1007/s12170-012-0222-8] [Medline: 23082235]
27. Kulick D. Wearable technology and long-term weight loss. J Am Med Assoc 2017 Jan 17;317(3):319. [doi: 10.1001/jama.2016.19256] [Medline: 28114541]
28. Jakicic JM, Belle SH. Wearable technology and long-term weight loss-reply. J Am Med Assoc 2017 Jan 17;317(3):319-320. [doi: 10.1001/jama.2016.19274] [Medline: 28114547]
29. Dyer O. Wearable fitness device does not help maintain weight loss, study finds. Br Med J 2016 Sep 25;354:i5204. [doi: 10.1136/bmj.i5204] [Medline: 27670652]
30. Assar S. PQDT Open - ProQuest. Ann Arbor, Michigan, USA: Dissertation Abstracts International; 2017. Evidence-Based Psychotherapeutic Interventions and mHealth for Weight Management in Overweight: A Biopsychosocial Framework URL: https://pqdtopen.proquest.com/doc/1875945106.html?FMT=AI [accessed 2020-02-05]
31. 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]
32. Booker CL, Harding S, Benzeval M. A systematic review of the effect of retention methods in population-based cohort studies. BMC Public Health 2011 Apr 19;11(1):249 [FREE Full text] [doi: 10.1186/1471-2458-11-249] [Medline: 21504610]
33. Canning KL, Brown RE, Jamnik VK, Salmon A, Ardern CI, Kuk JL. Individuals underestimate moderate and vigorous intensity physical activity. PLoS One 2014;9(5):e97927 [FREE Full text] [doi: 10.1371/journal.pone.0097927] [Medline: 24835105]
34. Kerr J, Marinac CR, Ellis K, Godbole S, Hipp A, Glanz K, et al. Comparison of accelerometry methods for estimating physical activity. Med Sci Sports Exerc 2017 Mar;49(3):617-624 [FREE Full text] [doi: 10.1249/MSS.0000000000001124] [Medline: 27755355]
35. Weiler DT, Villajuan SO, Edkins L, Cleary S, Saleem JJ. Wearable heart rate monitor technology accuracy in research: a comparative study between PPG and ECG technology. Proc Hum Factors Ergon Soc Annu Meet 2017;61(1):1292-1296. [doi: 10.1177/1541931213601804]
36. Coorey GM, Neubeck L, Mulley J, Redfern J. Effectiveness, acceptability and usefulness of mobile applications for cardiovascular disease self-management: Systematic review with meta-synthesis of quantitative and qualitative data. Eur J Prev Cardiol 2018 Mar;25(5):505-521. [doi: 10.1177/2047487317750913] [Medline: 29313363]
37. Brandt CJ, Clemensen J, Nielsen JB, Søndergaard J. Drivers for successful long-term lifestyle change, the role of e-health: a qualitative interview study. BMJ Open 2018 Mar 12;8(3):e017466 [FREE Full text] [doi: 10.1136/bmjopen-2017-017466] [Medline: 29530904]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 22.01.19; peer-reviewed by D Baines, K Ng, D Gratzer, M Bardus, T Fazzino, M Stuckey; comments to author 22.02.19; revised version received 06.11.19; accepted 16.12.19; published 10.03.20.

Please cite as:
Fawcett E, Van Velthoven MH, Meinert E
Long-Term Weight Management Using Wearable Technology in Overweight and Obese Adults: Systematic Review
JMIR Mhealth Uhealth 2020;8(3):e13461
URL: https://mhealth.jmir.org/2020/3/e13461
doi:10.2196/13461
PMID:32154788

Original Paper

Measurement of Step Angle for Quantifying the Gait Impairment of Parkinson's Disease by Wearable Sensors: Controlled Study

Jingying Wang[1,2], PhD; Dawei Gong[3], MM; Huichun Luo[1], PhD; Wenbin Zhang[4], MD; Lei Zhang[5], MSc; Han Zhang[1,2], MEng; Junhong Zhou[6*], PhD; Shouyan Wang[1,2*], PhD

[1]Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China

[2]Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (Fudan University), Ministry of Education, Shanghai, China

[3]Department of Neurosurgery, The Second Hospital of Nanjing, Nanjing, China

[4]Department of Neurosurgery, Nanjing Brain Hospital, Nanjing, China

[5]Department of Computer Science, Virginia Tech, Falls Church, VA, United States

[6]Hinda and Arthur Marcus Institute for Aging Research, Hebrew SeniorLife, Harvard Medical School, Roslindale, MA, United States

[*]these authors contributed equally

Corresponding Author:
Shouyan Wang, PhD
Institute of Science and Technology for Brain-Inspired Intelligence
Fudan University
220 Handan Road, Yangpu District
Shanghai, 200433
China
Phone: 86 181 1628 2313
Email: shouyan@fudan.edu.cn

Abstract

Background: Gait impairments including shuffling gait and hesitation are common in people with Parkinson's disease (PD), and have been linked to increased fall risk and freezing of gait. Nowadays the gait metrics mostly focus on the spatiotemporal characteristics of gait, but less is known of the angular characteristics of the gait, which may provide helpful information pertaining to the functional status and effects of the treatment in PD.

Objective: This study aimed to quantify the angles of steps during walking, and explore if this novel step angle metric is associated with the severity of PD and the effects of the treatment including the acute levodopa challenge test (ALCT) and deep brain stimulation (DBS).

Methods: A total of 18 participants with PD completed the walking test before and after the ALCT, and 25 participants with PD completed the test with the DBS on and off. The walking test was implemented under two conditions: walking normally at a preferred speed (single task) and walking while performing a cognitive serial subtraction task (dual task). A total of 17 age-matched participants without PD also completed this walking test. The angular velocity was measured using wearable sensors on each ankle, and three gait angular metrics were obtained, that is mean step angle, initial step angle, and last step angle. The conventional gait metrics (ie, step time and step number) were also calculated.

Results: The results showed that compared to the control, the following three step angle metrics were significantly smaller in those with PD: mean step angle ($F_{1,48}$=69.75, P<.001, partial eta-square=0.59), initial step angle ($F_{1,48}$=15.56, P<.001, partial eta-square=0.25), and last step angle ($F_{1,48}$=61.99, P<.001, partial eta-square=0.56). Within the PD cohort, both the ALCT and DBS induced greater mean step angles (ACLT: $F_{1,38}$=5.77, P=.02, partial eta-square=0.13; DBS: $F_{1,52}$=8.53, P=.005, partial eta-square=0.14) and last step angles (ACLT: $F_{1,38}$=10, P=.003, partial eta-square=0.21; DBS: $F_{1,52}$=4.96, P=.003, partial eta-square=0.09), but no significant changes were observed in step time and number after the treatments. Additionally, these step angles were correlated with the Unified Parkinson's Disease Rating Scale, Part III score: mean step angle (single task: r=–0.60, P<.001; dual task: r=–0.52, P<.001), initial step angle (single task: r=–0.35, P=.006; dual task: r=–0.35, P=.01), and last step angle (single task: r=–0.43, P=.001; dual task: r=–0.41, P=.002).

Conclusions: This pilot study demonstrated that the gait angular characteristics, as quantified by the step angles, were sensitive to the disease severity of PD and, more importantly, can capture the effects of treatments on the gait, while the traditional metrics

cannot. This indicates that these metrics may serve as novel markers to help the assessment of gait in those with PD as well as the rehabilitation of this vulnerable cohort.

(JMIR Mhealth Uhealth 2020;8(3):e16650) doi:10.2196/16650

KEYWORDS

Parkinson's disease; gait; angular velocity; inertial sensor; step angle; deep brain stimulation; acute levodopa challenge test

Introduction

Gait impairment, which is induced by diminished locomotor control [1], is highly prevalent in Parkinson's disease (PD) [2]. People with PD often suffer from multiple symptoms of gait impairment, including freezing of gait, hesitation at the beginning of walking, festination, and difficulty in stopping at the end of walking [3]. These gait impairments often lead to increased risk of falls, loss of functional independence in daily life, and even increased risk of morbidity and mortality [4]. Studies have linked the subtle changes in PD gait to other diseases and conditions such as dementia and history of head trauma [5]. It is thus of great clinical significance to measure and characterize the gait in PD, which will ultimately provide insights into the pathophysiology of PD and help optimize the therapeutic strategies such as deep brain stimulation (DBS).

Multiple instruments including motion capture systems [6], pressure mats [7], wearable sensors [8,9], and smartphone apps that use an accelerometer and gyroscope [10] have been developed to quantify the gait metrics of patients with PD [11]. Wearable sensors have become a rapidly growing solution to quantitatively assess the symptoms of PD, allowing more convenient testing protocol and remote and longer-term assessment and tracking of the functionality of people suffering from PD [12]. The wearable sensor can provide multiple spatiotemporal gait metrics of great clinical importance in PD assessment [12]. For example, the step time (or stride interval) [13,14] and step number (or step count) [13,15] can quantify the severity of locomotor dysfunction in those with PD and distinguish between the stages of PD. Additionally, gait speed [16-18], stride length [16,18], and cadence [17,18] can help identify abnormalities caused by PD and evaluate the efficacy of treatment.

Many impairments in PD, such as festinating gait, shuffling gait, hesitation, and start-stop difficulty have a visible resemblance (ie, the angles in the sagittal plane, the anatomical boundary dividing the left and right parts of the body, of the leg will shrink while walking). Several studies [19,20] observed significant correlation between the changes in the sagittal plane characteristics of the lower limbs and the clinical score, suggesting that such angular changes in the sagittal plane can provide meaningful clinical evaluations of PD. However, less is known about the angular characteristics of the sagittal plane gait in people with PD.

In this study, we aimed to explore if the angular characteristics of gait, especially in the sagittal plane, are sensitive to the clinical and functional characteristics of PD by measuring the step angles using wearable sensors fixed on the ankles. The mean step angle, step angle within the initiation of walking, and step angle within the end period of walking were measured. We hypothesize that the step angle would be significantly different between people with PD and those without PD (ie, control group); would be sensitive to the treatment (ie, medication and DBS) and cognitive demands; and would be significantly correlated to the Unified Parkinson's Disease Rating Scale, Part III (UPDRS-III) score.

Methods

Participants

A total of 30 participants with PD and 17 age-matched participants without PD were recruited. All participants provided written informed consent as approved by the institutional review committee of the Nanjing Brain Hospital. The inclusion criteria for the PD cohort were: having idiopathic PD as diagnosed by experienced clinicians based on the Chinese Diagnostic Criteria of Parkinson's Disease (2016), a surgery plan of DBS within 2 months, and able to stand and walk unassisted for more than 10 minutes. The exclusion criteria were: being younger than 40 years, having any other major neurological diseases (eg, stroke, dementia), and having ongoing psychiatric disturbances such as hallucinations.

Experimental Protocol

Before and after the treatments (ie, the acute levodopa challenge test [ALCT] or DBS), each participant in the PD group completed one 10-meter walking test for each of the following conditions: walking normally (ie, single task) and walking while performing a cognitive task (ie, dual task). The cognitive task was the serial subtraction of 3 or 7 from a random three-digit number. The UPDRS-III was also completed before and after the treatments and was used to assess the severity of PD. The healthy cohort completed one study visit consisting of the same single and dual task walking trials.

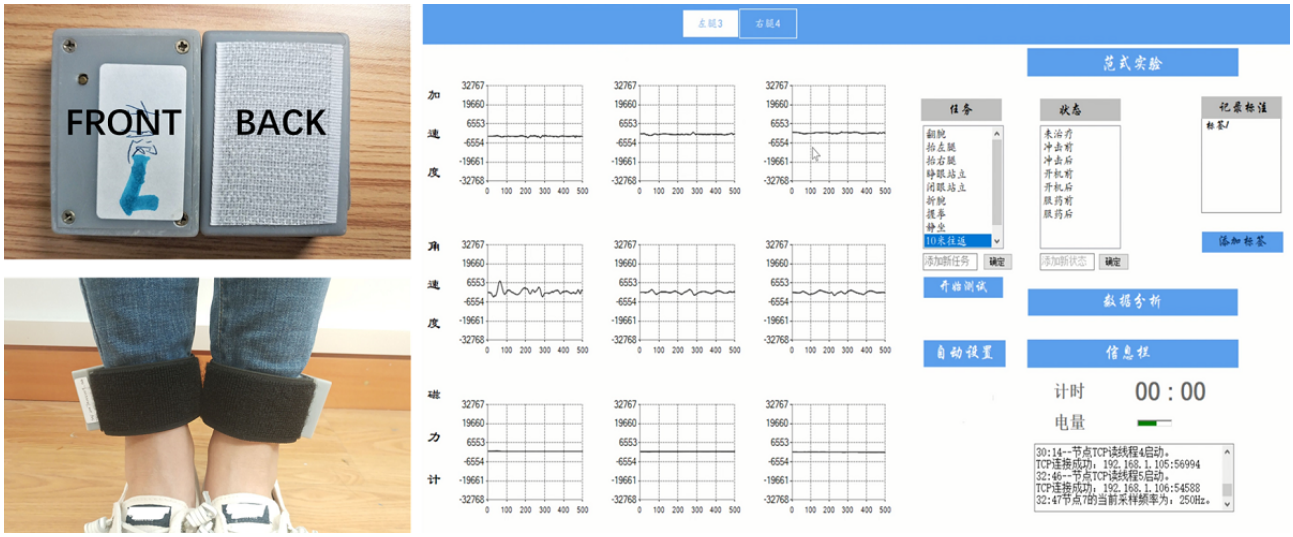
All trials were completed in the same room. One study personnel stood at the end of the walkway, and their position was fixed. No markers were used in the room, as the marker may give a cue to the participants in the PD group, and the pathway was not approaching a wall or doorway, which could interfere with participants' gait. Each participant stood in front of a wooden chair on one side of the pathway at the beginning of the trial and was instructed to walk along the pathway and stop at the same position as the study personnel. During each walking trial, two wearable sensors were used and attached to each ankle. The kinematic signals of walking including the angular velocity were then recorded and used to quantify the gait metrics.

Wearable Sensor

The wearable sensors (Figure 1) used in the study were developed by our team and embedded with inertial measurement

units (chip MPU-9250, IvenSense Inc, San Jose, CA). The sensors captured triaxial acceleration, angular velocity, and magnetic field intensity signals. The size of the sensors was 52 mm by 37 mm by 13 mm, and they each weighed 26.3 g. The average operating current of the sensors was 38.1 mA allowing for 13 hours of continuous recording. The sampling rate was 100 Hz, and the range of measurement for the gyroscope was ±1000 °/sec with a resolution of 0.06 °/sec/least significant bit.

Figure 1. Two sensors to measure the angular velocity of each ankle separately and the software interface of the sensors.



Acute Levodopa Challenge Test and Deep Brain Stimulation

The ALCT was performed in the morning using the established formulation to observe the clinical improvement in the PD group following withdrawal of all antiparkinsonian medication and overnight fasting [21]. The ALCT was used in clinics to screen patients who could use DBS based on their response to medication [21], that is if patients' UPDRS-III score reduced by more than 30% after the ALCT, they would benefit from DBS.

DBS is a type of neurosurgical surgery that sends electrical impulses through implanted electrodes to specific brain nuclei and alleviates the burden of multiple movement disorders such as tremors in PD [22]. In this study, the DBS surgery was conducted after a minimum 1-week break from the ALCT. The pulse of the DBS was set up using the width of 60 μs and a frequency of 130 Hz. On the DBS visit, which was during participants' perioperative period, participants completed the walking test with the DBS off and on without taking any medicine.

Data Processing

Figure 2 shows the pipeline of the data processing. First, we removed isolated noise points by using a 5-point median filter, and a band-pass filter was applied to remove fluctuations of frequency greater than 12 Hz and lower than 0.1 Hz. Second, the angular velocity of the axis with the largest variance was selected (Figure 3A). We observed that the angular changes in the sagittal plane during walking were obvious and much greater than the changes in other planes, so we focused on the angular characteristics of gait in this plane. Third, to divide gait cycles accurately, a threshold was set to exclude the influence of small displacements of the legs, and only the peaks and valleys that were greater than the threshold (ie, the heel strikes or toe-offs) were identified and used. The threshold was calculated as the mean of this signal plus or minus $\sqrt{2}$ / 2 SD (Figure 3B). The next step is the transformation from angular velocity to the degree of angle. The degree of angle was converted by the integral of the processed velocity signal. Finally, the step angle of each step was calculated as the difference of angle, $A_{(i)}$, between a peak of the wave, that is the angle of the maximum threshold=$T_{max(i)}$, and the adjacent valley, that is the angle of the minimum threshold=$T_{min(i)}$, in the angle degree series (Figure 3C).

Figure 2. The procedure of data processing.

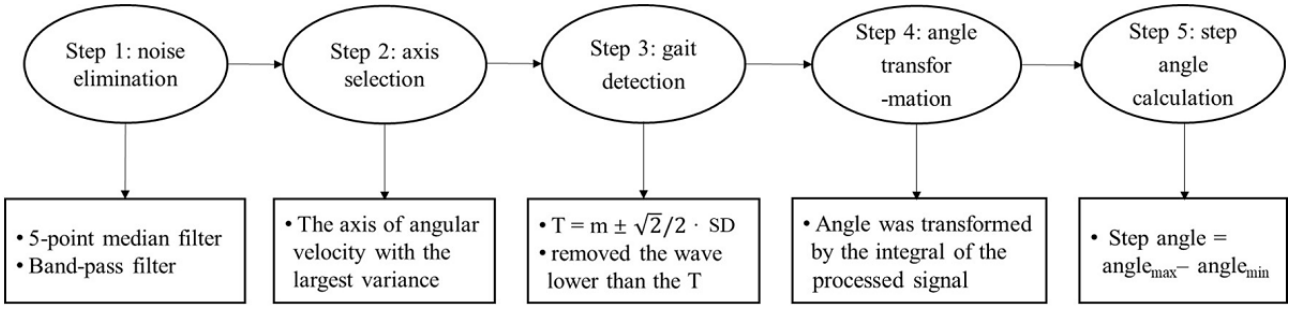
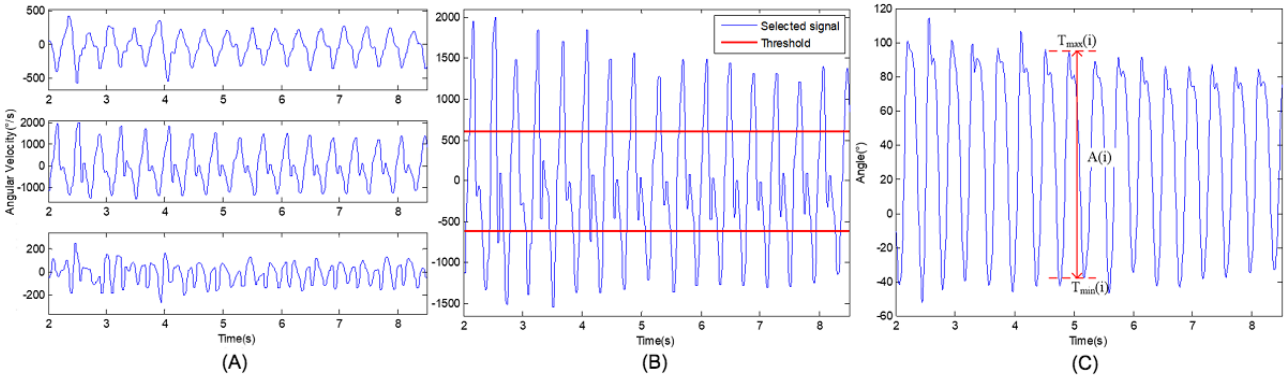


Figure 3. One signal in different steps of data processing: (A) 3-axis angular velocity signals; (B) the selected signal Y and thresholds; (C) the calculation of step angle.



Gait Metrics

Five gait metrics were calculated from the angular velocity signal, including 2 conventional gait metrics (step time and step number) and 3 new metrics (mean step angle, initial step angle, and last step angle). The definition of step angle is the angular change of the ankle in the sagittal plane within 1 gait cycle. The definitions of all 5 gait metrics are provided in Textbox 1. All the metrics here were obtained by averaging the left and right legs.

Textbox 1. Gait metrics and their definitions.

Conventional metrics

- Step time: the mean time to complete steps in 1 walking trial
- Step number: the total number of steps in 1 walking trial

New metrics

- Mean step angle: the arithmetic average of step angles in 1 walking trial
- Initial step angle: the angle of the first step in 1 walking trial
- Last step angle: the angle of the last step in 1 walking trail

Data Analysis

The statistical analysis was performed using SPSS 20 (IBM Corp, Armonk, NY). To examine the effects of group and task on step angle, 2-factor multivariate analyses of variance (MANOVAs) were used. The outcomes measured before the ALCT were used as the baseline outcomes of the PD group, as there was no influence for the implanted DBS electrodes. The dependent variables were mean step angle, initial step angle, and last step angle measured before the treatment (ie, baseline) in each model, and the model effects included the groups (ie, PD vs healthy), tasks (ie, single vs dual task), and their interactions. Similar models were also used to examine the effects of group and task on the two traditional metrics (ie, step time and number). To explore the effects of the ALCT on the UPDRS-III score and gait metrics, a 2-way MANOVA was used. The dependent variables were the 5 gait metrics and the UPDRS-III score. The model effects were time (before and after), task condition (single and dual task), and their interaction. The effects of DBS were then examined using the same models. If the interaction effect was significant, the analysis of simple effect was applied to further analyze the differences of the 5 gait metrics (dependent variables) between two levels of one independent variable while fixing another factor. The Bonferroni correction was used for multiple comparisons, where the significance level was set at $P<.01$. Partial eta-square was calculated for the effect size of the MANOVA. For partial eta-squares, 0.01, 0.06, and 0.14 were considered as small, medium, and large effect sizes, respectively [23].

Then a partial correlation analysis adjusted for age, gender, and status of treatment was used to explore the association between the UPDRS-III score and gait metrics at baseline. In addition, the independent-samples t test (two-tailed) and the Mann-Whitney U test were used to examine the differences in age and gender, respectively, between the two groups. A Bonferroni correction was used for the multiple comparisons, and the significance was set as $P<.01$.

The significance level was set at $P<.05$, not including those corrected by the Bonferroni correction.

Results

Demographics and Clinical Characteristics

A total of 18 participants in the PD group completed the tests at baseline (ie, before the ALCT), after the ALCT, and after the DBS surgery. In addition, 25 participants in the PD group completed the tests when the DBS was on and off. All participants completed the study. Table 1 shows their demographic and clinical information. No significant differences in age ($F_{2,57}=0.18$, $P=.84$) or gender ($\chi^2_2=1.66$, $P=.44$) between

the PD cohort and non-PD control group were observed. There were no significant differences of disease duration (t_{36}=0.16, *P*=.88), the Hoehn-Yahr stage (t_{32}=0.04, *P*=.97), and UPDRS-III score (t_{22}=–0.24, *P*=.81) between the ALCT and DBS groups.

Table 1. Demographics and clinical characteristics of participants.

Characteristics	Non-PD[a] (N=17)	PD with ALCT[b] treatment (N=18)	PD with DBS[c] treatment (N=25)
Gender, female, n (%)	8 (47%)	12 (67%)	16 (64%)
Age (years), mean (SD)	62.4 (7.1)	63.6 (5.9)	63.4 (6.9)
Disease duration (years), mean (SD)	n/a[d]	10.07 (2.65)	9.92 (3.08)
H-Y[e] stage, mean (SD)	n/a	3.29 (0.87)	3.28 (0.83)
UPDRS-III[f] score, mean (SD)	n/a	39.09 (13.33)	40.31 (11.65)

[a]PD: Parkinson's disease.

[b]ALCT: Acute Levodopa Challenge Test.

[c]DBS: Deep Brain Stimulation.

[d]Not applicable.

[e]H-Y: Hoehn-Yahr.

[f]UPDRS-III: Unified Parkinson's Disease Rating Scale, Part III.

Comparison of Gait Metrics Between People With Parkinson's Disease and Controls

A significant main effect of group was observed, but no significant difference was observed for the effect of task and their interaction (Table 2). The mean step angle ($F_{1,48}$=69.75, *P*<.001, partial eta-square=0.59), initial step angle ($F_{1,48}$=15.56, *P*<.001, partial eta-square=0.25), and last step angle ($F_{1,48}$=61.99, *P*<.001, partial eta-square=0.56) in participants with PD were all significantly smaller than healthy people. Similar results were shown in those conventional metrics that the PD cohort had significantly larger step times ($F_{1,48}$=7.52, *P*=.009, partial eta-square=0.14) and more step numbers ($F_{1,48}$=12.05, *P*=.001, partial eta-square=0.20) than the control.

Table 2. Gait metrics measured at baseline in participants with Parkinson's disease and healthy participants.

Metrics	non-Parkinson's disease, mean (SD)		Parkinson's disease, mean (SD)	
	Single	Dual	Single	Dual
Step time (sec/step)	0.89 (0.05)	0.97 (0.09)	1.04 (0.37)	1.2 (0.41)
Step number	30.62 (9.05)	34.34 (9.86)	50.95 (59.48)	89.94 (65.44)
Mean step angle (^{o})	63.33 (7.32)	61.85 (9.01)	30.71 (20.36)	26.87 (21.06)
Initial step angle (^{o})	34.50 (4.80)	32.88 (6.35)	33.71 (5.58)	20.79 (16.99)
Last step angle (^{o})	64.33 (15.20)	60.82 (14.21)	28.54 (15.63)	27.12 (16.56)

Effects of Acute Levodopa Challenge Test and Deep Brain Simulation on the Gait Metrics

We observed that the UPDRS-III score significantly decreased after the treatments (ie, the ALCT and DBS) compared to the score before the treatments (ACLT:t_{11}=–7.81, *P*<.001; DBS:t_{15}=–15.22, *P*<.001).

Acute Levodopa Challenge Test

A significant main effect of time (before the ALCT vs after the ALCT) was observed, but no significant main effect of task and their interaction were observed (Table 3). Specifically, the mean step angle ($F_{1,38}$=5.77, *P*=.02, partial eta-square=0.13) and last step angle ($F_{1,38}$=10, *P*=.003, partial eta-square=0.21) after the ALCT were greater than that at baseline, while the initial step angle ($F_{1,38}$=2.55, *P*=.12, partial eta-square=0.06) was not significantly changed. No significant changes were observed in the conventional metrics step number ($F_{1,38}$=4.33, *P*=.05, partial eta-square=0.10) and step time ($F_{1,38}$=2.01, *P*=.17, partial eta-square=0.05) after the ALCT.

Table 3. Gait metrics before and after acute levodopa challenge test.

Metrics	Before ALCT[a], mean (SD)		After ALCT, mean (SD)	
	Single task	Dual task	Single task	Dual task
Step time (sec/step)	1.04 (0.37)	1.20 (0.41)	0.96 (0.21)	1.01 (0.24)
Step number	50.95 (59.48)	89.94 (65.44)	32.21 (28.8)	46.77 (34.26)
Mean step angle (^{o})	30.71 (20.36)	26.87 (21.06)	46.48 (18.87)	40.63 (18.82)
Initial step angle (^{o})	20.79 (16.99)	21.26 (16.93)	32.99 (17.18)	26.03 (16.92)
Last step angle (^{o})	28.54 (15.63)	27.12 (16.56)	43.91 (17.37)	43.06 (13.66)

[a]ALCT: acute levodopa challenge test.

Deep Brain Stimulation

A significant main effect of time (DBS off vs DBS on) was observed, but not in the main effect of task and their interaction (Table 4). Specifically, mean step angle ($F_{1,52}$=8.53, P=.005, partial eta-square=0.14) and last step angle ($F_{1,52}$=4.96, P=.003, partial eta-square=0.09) with DBS on were both significantly greater than those with DBS off, but no significant changes in initial step angle ($F_{1,52}$=2.94, P=.09, partial eta-square=0.05) were observed. In conventional metrics, step time ($F_{1,52}$=5.59, P=.02, partial eta-square=0.1) had a marginally significant decrease when DBS was on, and no significant improvement was observed in step number ($F_{1,52}$=1.33, P=.25, partial eta-square=0.03).

Table 4. Gait metrics under deep brain stimulation off and deep brain stimulation on conditions.

Metrics	Deep brain stimulation off, mean (SD)		Deep brain stimulation on, mean (SD)	
	Single task	Dual task	Single task	Dual task
Step time (sec/step)	1.24 (0.65)	1.21 (0.48)	0.91 (0.23)	0.98 (0.28)
Step number	50.64 (82.6)	66.69 (41.31)	36.37 (23.01)	50.89 (23.35)
Mean step angle (^{o})	34.36 (17.77)	28.72 (17.52)	45.88 (16.61)	43.83 (16.25)
Initial step angle (^{o})	21.73 (12.59)	18.32 (9.55)	26.34 (12.93)	25.02 (13.61)
Last step angle (^{o})	33.59 (21.71)	25.47 (17.43)	40.55 (21.44)	42.53 (19.45)

Relationships Between Gait Metrics and Unified Parkinson's Disease Rating Scale, Part III

Table 5 presented the association between the gait metrics and the UPDRS-III score. Age, gender, and condition were controlled in this analysis, and the degree of freedom in single and dual walking tasks were 58 and 51, respectively. The mean, initial, and last step angles in both single and dual walking tasks were all significantly correlated with the UPDRS-III score (r>–0.35, P<.01). A weaker correlation was observed between the step time and number and the UPDRS-III scores. The scatter plots of the UPDRS-III score and gait metrics are shown in Multimedia Appendix 1.

Table 5. Partial correlation analysis between the Unified Parkinson's Disease Rating Scale, Part III score and gait metrics.

Metrics	Single task, *r*	*P* value	Dual task, *r*	*P* value
Step time (sec/step)	0.32	.01	0.42	.002
Step number	0.18	.17	0.35	.01
Mean step angle (^{o})	–0.60	<.001	–0.52	<.001
Initial step angle (^{o})	–0.35	.006	–0.35	.01
Last step angle (^{o})	–0.43	.001	–0.41	.002

Discussion

This study demonstrated that the angular characteristics of gait, as quantified using the step angles measured in the sagittal plane of the lower limbs, are sensitive to PD and the treatments (ie, the ALCT and DBS) and consistent in different cognitive conditions. Specifically, we observed that the mean, initial, and last step angle were significantly smaller in the PD cohort compared to the healthy cohort and similar between single and dual task conditions. In addition, the mean step angle and last step angle were significantly increased after treatments and were associated with the UPDRS-III score. These results suggest that these novel angular metrics are sensitive to the severity of PD

and captures the effects of treatments on gait in people with PD, as greater step angles reflected better locomotor control of walking.

We observed that the new angular metrics were smaller in those with PD across the conditions (single and dual task) compared to the healthy cohort and significantly correlated with the UPDRS-III total score. Participants with greater UPDRS-III scores had smaller step angles. The commonly used gait metrics focused more on the temporal (eg, stride time) or spatial (eg, stride length) characteristics of gait, but the musculoskeletal rotation of the extremities is also important for the completion of 1 gait cycle. The step angle captures the angular change (ie, rotation) in the sagittal plane of lower limbs during walking. The diminished locomotor control in PD may induce more variance in the rotation and thus impair gait patterns (eg, incomplete gait cycles). These step angles measuring the subtle changes in the musculoskeletal rotation may thus help quantify the gait impairments in PD. The initial and last step angles, for example, can particularly help assess the start hesitation and stop difficulty in PD. It should also be noted that compared to those traditional metrics (step time, step number), which have been proved effective in reflecting gait impairments in patients with PD [13], the effect size of the angular metrics is much larger, indicating that these new metrics are more sensitive to the effects of PD on gait. Future longitudinal studies are needed to explore how these angular characteristics of gait change along with the progress of PD.

These step angles were sensitive to both the ALCT and DBS within the PD cohort. This is consistent with the results of previous studies using other gait metrics [24]. Specifically, the mean and last step angles were increased when DBS was working, and the mean and last step angles were improved after the ALCT. However, no significant improvement was observed in those traditional metrics. These results support that the subthalamic nucleus (STN)-DBS treatment improves gait performance in PD [25-27]; however, some studies [28,29] reported the effects of STN-DBS on gait are less successful and may even lead to an aggravation of freezing of gait and imbalance. The results here suggested that step angle metrics may capture the subtle changes of gait and the acute effects of the treatment on gait. The effect size of the ALCT was larger than that of DBS, which is in support of previous studies showing a lower benefit on gait velocity and stride length by DBS [30] and a higher benefit on cadence by levodopa [31]. However, this still suggests that the STN-DBS has less effectiveness on gait compared to the levodopa treatment.

Previous studies have shown that the performance of simultaneous cognitive tasks compromised gait in people with and without PD [32,33] with a decrease in stride time and an increase in stride frequency [34]. However, this study showed that compared to single-task walking, no statistically significant change in step time, step number, and step angles were observed in dual task walking. One potential reason may be the relatively small sample size in this piloting study. However, a greater change in step angle from single to dual task walking in those with PD (average change of 5°) compared to the control cohort (smaller difference of only 2°) was observed, indicating that PD diminished the capacity of walking control in the dual task condition.

The results of this study further provided evidence that wearable inertial sensors can help advance the traditional measurements of gait and other neurophysiological and biomechanical signals (eg, the center of pressure of a human body) into a quicker, convenient protocol [35]. Traditional laboratory or clinical tests assessing the gait are dependent upon expensive and nonportable equipment and well-trained study personnel, which presents a challenge for people living in rural areas or areas distant from hospitals. This type of wearable sensor provides a novel approach to these populations for assessing their gait and other health information. It should be noted that the intersubject variance of gait metrics was much greater (ie, larger SD) in the PD cohort compared to the healthy cohort in this study, indicating that subtle characteristics of gait vary across the PD population. Previous studies also showed that within one person, the day-to-day variance in gait and physical activity was high, and such variance was associated with their health status, such as cognitive impairments [36]. Taken together, the measurement of gait and other health information using wearable sensors facilitates the high-frequency monitoring in those vulnerable populations, which will ultimately help in clinical diagnosis and disease prevention.

The small sample size and lack of cognition examination are two limitations of this study. The difference in cognitive function, which may influence the gait, between the two groups was not included in the analyses. Future studies with a larger sample size and assessment of cognitive function, using the Montreal Cognitive Assessment or the Mini Mental State Examination, are needed to examine and confirm the findings in this pilot study. The flexibility of the knee also contributes to the control of musculoskeletal rotation of gait, and thus the angular change of the knee may also provide important information about locomotor control. Future work using an electronic goniometer is needed to measure the angular characteristics of the knee during walking. Nevertheless, this study proposed novel metrics to quantify the angular characteristics of gait and demonstrated that the step angle metrics are sensitive to the effects of PD on gait, disease severity, and the effects of treatments on gait, which may serve as novel markers that help the management of PD.

Acknowledgments

This research was supported by the Chinese Ministry of Education 111 Project (No B18015); the Key Project of Shanghai Municipal Commission of Science and Technology (No 16JC1420402); Shanghai municipal commission of science and technology Major Project (No 2018SHZDZX01) and ZJ Lab; National Key R&D Program of China (No 2018YFC1312900); Shanghai Municipal Commission of Science and Technology (No 17JC1401400); National Key R&D Program of China (No

2018YFC1705800); Social Development of Jiangsu Province Key Research Plan (No BE2016614); Special Project of Ministry of Science and Technology on Digital Medical Equipment (No 2016YFC0105900); Research BWS16J011 and BWS16J035.

Dr Junhong Zhou is supported by Fudan Scholar program and Hebrew SeniorLife Applebaum grant.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Scatter plots of the UPDRS-III score and gait metrics.
[PDF File (Adobe PDF File), 302 KB - mhealth_v8i3e16650_app1.pdf]

References

1. Earhart GM. Dynamic control of posture across locomotor tasks. Mov Disord 2013 Sep 15;28(11):1501-1508 [FREE Full text] [doi: 10.1002/mds.25592] [Medline: 24132838]
2. Stolze H, Klebe S, Baecker C, Zechlin C, Friege L, Pohle S, et al. Prevalence of gait disorders in hospitalized neurological patients. Mov Disord 2005 Jan 18;20(1):89-94. [doi: 10.1002/mds.20266] [Medline: 15390043]
3. Tan D, Danoudis M, McGinley J, Morris ME. Relationships between motor aspects of gait impairments and activity limitations in people with Parkinson's disease: a systematic review. Parkinsonism Relat Disord 2012 Feb;18(2):117-124. [doi: 10.1016/j.parkreldis.2011.07.014] [Medline: 22093237]
4. Boonstra TA, van der Kooij H, Munneke M, Bloem BR. Gait disorders and balance disturbances in Parkinson s disease: clinical update and pathophysiology. Current Opinion in Neurology 2008;24(4):461-471. [doi: 10.1097/wco.0b013e328305bdaf]
5. Manor B, Zhou J, Lo O, Zhu H, Gouskova NA, Yu W, et al. Self-reported head trauma predicts poor dual task gait in retired National Football League players. Ann Neurol 2020 Jan 29;87(1):75-83 [FREE Full text] [doi: 10.1002/ana.25638] [Medline: 31693765]
6. Defebvre L, Blatt J, Blond S, Bourriez J, Gieu JD, Destée A. Effect of thalamic stimulation on gait in Parkinson disease. Arch Neurol 1996 Sep 01;53(9):898-903. [doi: 10.1001/archneur.1996.00550090108016] [Medline: 8815855]
7. Okuno R, Fujimoto S, Akazawa J, Yokoe M, Sakoda S, Akazawa K. Analysis of spatial temporal plantar pressure pattern during gait in Parkinson's disease. 2008 Aug 24 Presented at: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2008; Vancouver p. 1765-1768. [doi: 10.1109/iembs.2008.4649519]
8. Mariani B, Jiménez MC, Vingerhoets FJG, Aminian K. On-shoe wearable sensors for gait and turning assessment of patients with Parkinson's disease. IEEE Trans Biomed Eng 2013 Jan;60(1):155-158. [doi: 10.1109/tbme.2012.2227317]
9. Stack E, Agarwal V, King R, Burnett M, Tahavori F, Janko B, et al. Identifying balance impairments in people with Parkinson's disease using video and wearable sensors. Gait Posture 2018 May;62:321-326. [doi: 10.1016/j.gaitpost.2018.03.047] [Medline: 29614464]
10. Mazilu S, Hardegger M, Zhu Z, Roggen D, Tröster G, Plotnik M, et al. Online detection of freezing of gait with smartphones and machine learning techniques. 2012 May Presented at: 6th International Conference on Pervasive Computing Technologies for Healthcare; 2012; San Diego p. 123-130. [doi: 10.4108/icst.pervasivehealth.2012.248680]
11. Espay AJ, Bonato P, Nahab FB, Maetzler W, Dean JM, Klucken J, Movement Disorders Society Task Force on Technology. Technology in Parkinson's disease: challenges and opportunities. Mov Disord 2016 Sep;31(9):1272-1282 [FREE Full text] [doi: 10.1002/mds.26642] [Medline: 27125836]
12. Maetzler W, Domingos J, Srulijes K, Ferreira JJ, Bloem BR. Quantitative wearable sensors for objective assessment of Parkinson's disease. Mov Disord 2013 Oct 12;28(12):1628-1637. [doi: 10.1002/mds.25628] [Medline: 24030855]
13. Keloth SM, Viswanathan R, Jelfs B, Arjunan S, Raghav S, Kumar D. Gait parameters and walking patterns show the significant differences between Parkinson's disease and healthy participants? Biosensors (Basel) 2019 Apr 25;9(2) [FREE Full text] [doi: 10.3390/bios9020059] [Medline: 31027153]
14. Bartsch R, Plotnik M, Kantelhardt JW, Havlin S, Giladi N, Hausdorff JM. Fluctuation and synchronization of gait intervals and gait force profiles distinguish stages of Parkinson's disease. Physica A 2007 Sep 15;383(2):455-465 [FREE Full text] [doi: 10.1016/j.physa.2007.04.120] [Medline: 18163154]
15. King LA, Mancini M, Priest K, Salarian A, Rodrigues-de-Paula F, Horak F. Do clinical scales of balance reflect turning abnormalities in people with Parkinson's disease? J Neurol Phys Ther 2012 Mar;36(1):25-31 [FREE Full text] [doi: 10.1097/NPT.0b013e31824620d1] [Medline: 22333919]
16. Moore ST, MacDougall HG, Gracies J, Ondo WG. Locomotor response to levodopa in fluctuating Parkinson's disease. Exp Brain Res 2008 Feb 8;184(4):469-478. [doi: 10.1007/s00221-007-1113-y] [Medline: 17828529]
17. Zampieri C, Salarian A, Carlson-Kuhta P, Aminian K, Nutt JG, Horak FB. The instrumented timed up and go test: potential outcome measure for disease modifying therapies in Parkinson's disease. J Neurol Neurosurg Psychiatry 2010 Feb;81(2):171-176 [FREE Full text] [doi: 10.1136/jnnp.2009.173740] [Medline: 19726406]

18. Schlachetzki JCM, Barth J, Marxreiter F, Gossler J, Kohl Z, Reinfelder S, et al. Wearable sensors objectively measure gait parameters in Parkinson's disease. PLoS One 2017 Oct 11;12(10):e0183989 [FREE Full text] [doi: 10.1371/journal.pone.0183989] [Medline: 29020012]
19. Morris M, Iansek R, McGinley J, Matyas T, Huxham F. Three-dimensional gait biomechanics in Parkinson's disease: evidence for a centrally mediated amplitude regulation disorder. Mov Disord 2005 Jan 13;20(1):40-50. [doi: 10.1002/mds.20278] [Medline: 15390033]
20. Grasso R, Peppe A, Stratta F, Angelini D, Zago M, Stanzione P, et al. Basal ganglia and gait control: apomorphine administration and internal pallidum stimulation in Parkinson's disease. Exp Brain Res 1999 May 4;126(2):139-148. [doi: 10.1007/s002210050724] [Medline: 10369137]
21. Albanese A, Bonuccelli U, Brefel C, Chaudhuri KR, Colosimo C, Eichhorn T, et al. Consensus statement on the role of acute dopaminergic challenge in Parkinson's disease. Mov Disord 2001 Mar;16(2):197-201. [doi: 10.1002/mds.1069]
22. Benabid AL. Deep brain stimulation for Parkinson's disease. Curr Opin Neurobiol 2003 Dec;13(6):696-706. [doi: 10.1016/j.conb.2003.11.001] [Medline: 14662371]
23. Cohen J. Statistical Power Analyses for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
24. McNeely ME, Earhart GM. Medication and subthalamic nucleus deep brain stimulation similarly improve balance and complex gait in Parkinson disease. Parkinsonism Relat Disord 2013 Jan;19(1):86-91. [doi: 10.1016/j.parkreldis.2012.07.013] [Medline: 22885253]
25. McNeely ME, Hershey T, Campbell MC, Tabbal SD, Karimi M, Hartlein JM, et al. Effects of deep brain stimulation of dorsal versus ventral subthalamic nucleus regions on gait and balance in Parkinson's disease. J Neurol Neurosurg Psychiatry 2011 Nov;82(11):1250-1255 [FREE Full text] [doi: 10.1136/jnnp.2010.232900] [Medline: 21478202]
26. Mera TO, Filipkowski DE, Riley DE, Whitney CM, Walter BL, Gunzler SA, et al. Quantitative analysis of gait and balance response to deep brain stimulation in Parkinson's disease. Gait Posture 2013 May;38(1):109-114 [FREE Full text] [doi: 10.1016/j.gaitpost.2012.10.025] [Medline: 23218768]
27. Hausdorff JM, Gruendlinger L, Scollins L, O'Herron S, Tarsy D. Deep brain stimulation effects on gait variability in Parkinson's disease. Mov Disord 2009 Aug 15;24(11):1688-1692. [doi: 10.1002/mds.22554] [Medline: 19554569]
28. Tommasi G, Lopiano L, Zibetti M, Cinquepalmi A, Fronda C, Bergamasco B, et al. Freezing and hypokinesia of gait induced by stimulation of the subthalamic region. J Neurol Sci 2007 Jul 15;258(1-2):99-103. [doi: 10.1016/j.jns.2007.03.002] [Medline: 17445832]
29. Stolze H, Klebe S, Poepping M, Lorenz D, Herzog J, Hamel W, et al. Effects of bilateral subthalamic nucleus stimulation on parkinsonian gait. Neurology 2001 Jul 10;57(1):144-146. [doi: 10.1212/wnl.57.1.144] [Medline: 11445647]
30. Krystkowiak P, Blatt J, Bourriez J, Duhamel A, Perina M, Blond S, et al. Effects of subthalamic nucleus stimulation and levodopa treatment on gait abnormalities in Parkinson disease. Arch Neurol 2003 Jan 01;60(1):80. [doi: 10.1001/archneur.60.1.80]
31. Ferrarin M, Rizzone M, Lopiano L, Recalcati M, Pedotti A. Effects of subthalamic nucleus stimulation and l-dopa in trunk kinematics of patients with Parkinson's disease. Gait & Posture 2004 Apr;19(2):164-171. [doi: 10.1016/s0966-6362(03)00058-4]
32. O'Shea S, Morris M, Iansek R. Dual task interference during gait in people with parkinson disease: effects of motor versus cognitive secondary tasks. Physical Therapy 2002 Sep 1;82(9):888-897. [doi: 10.1093/ptj/82.9.888]
33. Amboni M, Barone P, Iuppariello L, Lista I, Tranfaglia R, Fasano A, et al. Gait patterns in Parkinsonian patients with or without mild cognitive impairment. Mov Disord 2012 Oct 02;27(12):1536-1543. [doi: 10.1002/mds.25165] [Medline: 23032876]
34. Ebersbach G, Dimitrijevic MR, Poewe W. Influence of concurrent tasks on gait: a dual-task approach. Percept Mot Skills 1995 Aug 04;81(1):107-113. [doi: 10.2466/pms.1995.81.1.107] [Medline: 8532444]
35. Mancini M, King L, Salarian A, Holmstrom L, McNames J, Horak FB. Mobility lab to assess balance and gait with synchronized body-worn sensors. J Bioengineer & Biomedical Sci 2012. [doi: 10.4172/2155-9538.s1-007]
36. Hayes TL, Abendroth F, Adami A, Pavel M, Zitzelberger TA, Kaye JA. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. Alzheimers Dement 2008 Nov 01;4(6):395-405 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.004] [Medline: 19012864]

Abbreviations

ALCT: acute levodopa challenge test
DBS: deep brain stimulation
MANOVA: multivariate analysis of variance
PD: Parkinson’s disease
STN: subthalamic nucleus
UPDRS-III: Unified Parkinson's Disease Rating Scale, Part III.

Edited by G Eysenbach; submitted 12.10.19; peer-reviewed by B Price, T Georgiou, B Cole; comments to author 07.11.19; revised version received 29.12.19; accepted 07.02.20; published 20.03.20.

Please cite as:
Wang J, Gong D, Luo H, Zhang W, Zhang L, Zhang H, Zhou J, Wang S
Measurement of Step Angle for Quantifying the Gait Impairment of Parkinson's Disease by Wearable Sensors: Controlled Study
JMIR Mhealth Uhealth 2020;8(3):e16650
URL: http://mhealth.jmir.org/2020/3/e16650/
doi:10.2196/16650
PMID:32196458

Original Paper

Accurate Measurement of Handwash Quality Using Sensor Armbands: Instrument Validation Study

Chaofan Wang[1], BSc, MSc; Zhanna Sarsenbayeva[1], BSc, MSc; Xiuge Chen[1], BSc; Tilman Dingler[1], BSc, MSc, PhD; Jorge Goncalves[1], BSc, MSc, PhD; Vassilis Kostakos[1], BSc, PhD

School of Computing and Information Systems, The University of Melbourne, Parkville, Australia

Corresponding Author:
Chaofan Wang, BSc, MSc
School of Computing and Information Systems
The University of Melbourne
Doug McDonell Building (Building 168)
Parkville, 3010
Australia
Phone: 61 3903 58966
Email: chaofan.wang@unimelb.edu.au

Abstract

Background: Hand hygiene is a crucial and cost-effective method to prevent health care–associated infections, and in 2009, the World Health Organization (WHO) issued guidelines to encourage and standardize hand hygiene procedures. However, a common challenge in health care settings is low adherence, leading to low handwashing quality. Recent advances in machine learning and wearable sensing have made it possible to accurately measure handwashing quality for the purposes of training, feedback, or accreditation.

Objective: We measured the accuracy of a sensor armband (Myo armband) in detecting the steps and duration of the WHO procedures for handwashing and handrubbing.

Methods: We recruited 20 participants (10 females; mean age 26.5 years, SD 3.3). In a semistructured environment, we collected armband data (acceleration, gyroscope, orientation, and surface electromyography data) and video data from each participant during 15 handrub and 15 handwash sessions. We evaluated the detection accuracy for different armband placements, sensor configurations, user-dependent vs user-independent models, and the use of bootstrapping.

Results: Using a single armband, the accuracy was 96% (SD 0.01) for the user-dependent model and 82% (SD 0.08) for the user-independent model. This increased when using two armbands to 97% (SD 0.01) and 91% (SD 0.04), respectively. Performance increased when the armband was placed on the forearm (user dependent: 97%, SD 0.01; and user independent: 91%, SD 0.04) and decreased when placed on the arm (user dependent: 96%, SD 0.01; and user independent: 80%, SD 0.06). In terms of bootstrapping, user-dependent models can achieve more than 80% accuracy after six training sessions and 90% with 16 sessions. Finally, we found that the combination of accelerometer and gyroscope minimizes power consumption and cost while maximizing performance.

Conclusions: A sensor armband can be used to measure hand hygiene quality relatively accurately, in terms of both handwashing and handrubbing. The performance is acceptable using a single armband worn in the upper arm but can substantially improve by placing the armband on the forearm or by using two armbands.

(JMIR Mhealth Uhealth 2020;8(3):e17001) doi:10.2196/17001

KEYWORDS

hand hygiene; wearable devices; machine learning

Introduction

Background

Health care workers' (HCWs') hands play a pivotal role in spreading microorganisms in health care environments and pose a direct clinical threat to patients [1-3]. Health care–associated infections (HAIs), also known as *nosocomial* infections, are the most common causes of morbidity and mortality in hospitals around the world [4]. The average prevalence of HAIs varies from 4.0% to 15.5% in different countries and regions [5-8].

De Angelis et al [9] estimated that patients with HAIs spent approximately 2.5 times more time in hospital and incurred costs that were 2.8 times higher than those for patients free from infection. Moreover, HAIs put a considerable financial burden on the health system: in France alone, nearly 3% of surgical procedures performed in 2010 resulted in infections creating annual costs of €58 million (US $64.5 million) [10], whereas in the United States, the estimated annual costs range from US $28 billion to US $45 billion [11]. Hand hygiene is a simple and cost-effective intervention to prevent HAIs and reduce their transmission [1,12].

In 2009, the World Health Organization (WHO) issued guidelines on Hand Hygiene in Health Care to provide a thorough review of evidence on hand hygiene and recommendations to promote hand hygiene in health care environments [3,13]. The guidelines summarize the key moments for hand hygiene for HCWs as *My 5 Moments For Hand Hygiene*, which have since been the focus of much work on automated technologies [14]. Previous work has reported that the compliance rate of hand hygiene is unacceptably poor [15], with a meta-analysis of 96 empirical studies showing a median compliance rate of 40% among HCWs [16]. A variety of technologies have been developed to increase compliance with the *5 moments for hand hygiene*, such as monitoring staff movement [17], using radio-frequency identification or Bluetooth beacons [14,18,19], and a range of commercial products, such as BioVigil (BioVigil Healthcare Systems, Inc) and SwipeSense (SwipeSense, Inc).

However, the WHO guidelines also introduced procedures for alcohol-based handrub and handwash with soap and water (shown in Figure 1 [3,20]). These guidelines aim to improve hand hygiene by decreasing colonization with transient flora [12,13]. Especially, alcohol-based handrub is preferred for routine decontamination of hands for all clinical indications, whereas handwash with soap and water is recommended for visibly soiled hands [3]. Different from the compliance rate of hand hygiene, which focuses on whether HCWs perform hand hygiene on time, the compliance with the WHO hand hygiene routines ensures that HCWs achieve adequate coverage of all surfaces of their hands with hand hygiene products [3].

Owing to the technical challenge of monitoring compliance with this set of guidelines, previous work measured compliance with this set of guidelines based on direct observation by trained auditors [21-23]. Yet, compliance with this set of WHO guidelines has been reported to be approximately 8.5% [23]. Crucially, in a large-scale assessment of hand hygiene quality conducted in Singapore, only 72% of HCWs could achieve adequate coverage of all hand surfaces *immediately* after hand hygiene training [24]. These findings suggest that it is important to develop reliable instruments for monitoring the quality of hand hygiene in health care settings by measuring adherence to the WHO handrub and handwash procedures.

Although a wide range of technologies have been developed to measure hand hygiene compliance rate, this is not the case for hand hygiene quality. Most work on quality monitoring has so far relied on cameras. Llorca et al [25] mounted a red green blue camera above a sink to collect hand hygiene practices among HCWs. After performing analysis on color and motion, their system used support vector machine (SVM) to classify six steps of a standard hand hygiene procedure. Similarly, Xia et al [26] collected hand hygiene videos through a red green blue and depth camera. By applying SVM, random forest, and linear discriminant analysis to the collected videos, their system could determine 12 steps of a hand hygiene procedure at the level of a single frame or a single video. A commercially available system is SureWash (Glanta Ltd), which can detect HCWs' hand motion and provide reminders for the upcoming steps, according to the standard WHO hand hygiene procedure. There has also been some work on hand hygiene–monitoring systems that provide assistance for people with dementia [27]. However, a major concern for camera-based systems to monitor hand hygiene quality is privacy, as those systems inevitably require the installation of cameras in toilets and patient care areas. Furthermore, cameras can only be used in certain environments, and camera-based monitoring systems may not provide actionable data on all hand hygiene events [28]. An additional challenge is that installed cameras cannot be easily moved should the need arise. Especially, cameras are not suitable for monitoring handrub because it tends to happen on the move, although a camera has a fixed field of view. Another concern is that the costs associated with the installation and maintenance of camera systems can be substantial [14].

An alternative to using camera-based systems is to use wearable sensors [29-32]. A common practice in literature is through the use of wristbands [30-32]. Yet, because of hygiene concerns, the WHO recommends removing rings, wristwatches, and bracelets before beginning surgical hand preparation [3,13]. Therefore, it is challenging to use wristbands to monitor hand hygiene procedure compliance, as it can possibly defeat the purpose of hand hygiene. Another limitation with previous studies is the use of Hidden Markov Model (HMM) to classify the steps of the hand hygiene procedure from feature vectors [30] or smooth classification results [29], which assumes HCWs will perform hand hygiene procedures according to the predefined orders. However, once that assumption is relaxed, the performance of these systems dramatically drops, for example, from 85% to 69% [30]. Similarly, previous studies make assumptions about the duration of hand hygiene steps and, for example, collect each step of the hand hygiene procedure individually with, say, a 5-second window [29]. Nonetheless, hand hygiene steps can actually vary rather substantially in duration, from 20 to 30 seconds for handwash to 40 to 60 seconds for handrub [3].

Figure 1. Standard World Health Organization procedures of alcohol-based handrub and handwash with soap and water. Source: World Health Organization. How to Handrub? / How to Handwash? [20]

Objectives

In this paper, we evaluated the accuracy of a sensor armband in measuring compliance with the WHO handrub and handwash guidelines. We used the Myo armband (North Inc) worn on participants' forearm or arm and evaluated a machine learning classifier that uses eXtreme Gradient Boosting (XGBoost) and E.Divisive [33] to identify each step of the standard WHO handwash or handrub procedures. The analysis evaluated eight different armband placements, six sensor configurations, user-dependent vs user-independent models, and bootstrapping performance.

The contributions of this paper are as follows: (1) unified model—although most previous work focuses on either handrub or handwash [30,34], our work combines these two common activities to create a unified hand hygiene model. Our model can detect each step of the respective WHO procedures. (2) Flexible detection—previous work assumes that HCWs perform hand hygiene according to the expected step order and duration [29,30]. Our model can detect each step individually, regardless of its order, timing, or duration, and can thus provide granular and detailed feedback. (3) Placement recommendations—by analyzing the rich data collected in our study, we can quantify the trade-off between classification accuracy and sensor placement and provide placement recommendations. (4) Sensor recommendations—as the Myo armband consists of a nine-axis inertial measurement unit (IMU) and eight electromyographic electrodes, we were able to collect acceleration, gyroscope, orientation, and surface electromyography (sEMG) data

simultaneously [35]. By evaluating different sensor data combinations, we were able to quantify the trade-off between accuracy and sensor types and recommend sensors for future hand hygiene–monitoring systems.

Methods

Task and Hardware Specification

We adopted the procedures for alcohol-based handrub and handwash with soap and water (Figure 1 [20]) recommended by the WHO guidelines on Hand Hygiene in Health Care [3]. The procedure contains seven steps for handrub and 11 steps for handwash (shown in Figure 1 [20]). Some of the steps are repeated for each hand. The guidelines claim that the total duration of the handrub routine is 20 to 30 seconds on average, whereas handwash lasts for 40 to 60 seconds.

We summarized all the steps for both procedures in Table 1. For handwash, we decided to exclude step 0 (wet hands with water) and step 1 (apply hand hygiene product), as the position of faucets and hand hygiene dispensers can vary across different lavatories. Furthermore, as previous work has noted [34], hand dominance is not crucial to hand hygiene detection because arm movements are symmetric. For this reason, our analysis does not consider the hand dominance effect.

Table 1. Steps of the World Health Organization hand hygiene procedures.

Step #	Steps of the World Health Organization hand hygiene procedures	
	Alcohol-based handrub	Handwash with soap and water
0	N/A[a]	N/A
1	Apply hand hygiene products	N/A
2	Rub palm	Rub palm
3 (R[b])	Rub dorsum (R)	Rub dorsum (R)
3 (L[c])	Rub dorsum (L)	Rub dorsum (L)
4	Interlock fingers	Interlock fingers
5 (R)	Twist knuckles (R)	Twist knuckles (R)
5 (L)	Twist knuckles (L)	Twist knuckles (L)
6 (R)	Rub thumb (R)	Rub thumb (R)
6 (L)	Rub thumb (L)	Rub thumb (L)
7 (R)	Scrub fingertip (R)	Scrub fingertip (R)
7 (L)	Scrub fingertip (L)	Scrub fingertip (L)
8	N/A	Rinse hands
9	N/A	Dry hands with towel
10	N/A	Turn off faucet with towel

[a]Not available.

[b]R: right.

[c]L: left.

The hardware for this study was the Myo armband, which contains a nine-axis IMU sensor and eight electromyographic electrodes. Through the nine-axis IMU sensor, the armband captures acceleration, gyroscope, and orientation data at a sample rate of 50 Hz. Using the eight surface electromyographic electrodes, it can also collect sEMG signals to measure users' muscular activity at a frequency of 200 Hz. Physiological data were transmitted to a receiver (in our case, a laptop) via the Bluetooth Low Energy protocol. Data from multiple armbands were simultaneously collected using the Myo SDK [35] and modified myo-python library [36] and synchronized using the algorithm proposed by Wang et al [37].

Experiment Design

Our experimental design had four independent variables: hygiene mode (handwash vs handrub), instruction mode (video vs poster), armband placement (above elbow vs below elbow), and hand (left vs right hand). We measured the time taken to complete each step of each hygiene procedure and the errors made by participants when washing their hands. We also measured the accuracy of the classifier in detecting each step of the procedure. The experiment followed a within-subjects design, and all participants completed all conditions in a counterbalanced manner. The University of Melbourne's Engineering Human Ethics Advisory Group approved the study.

We recruited 20 participants through our university's mailing lists and snowball recruitment with an equal number of women and men. All participants were students or staff in our university, and their ages ranged between 22 and 33 years (mean 26.5 years, SD 3.31). The majority of participants (18/20, 90%) had not received formal training in hand hygiene in the last 3 years and were not familiar with the formal hand hygiene procedures. A

total of 95% (19/20) of participants reported using their right hand as their dominant hand and no ambidextrous participants.

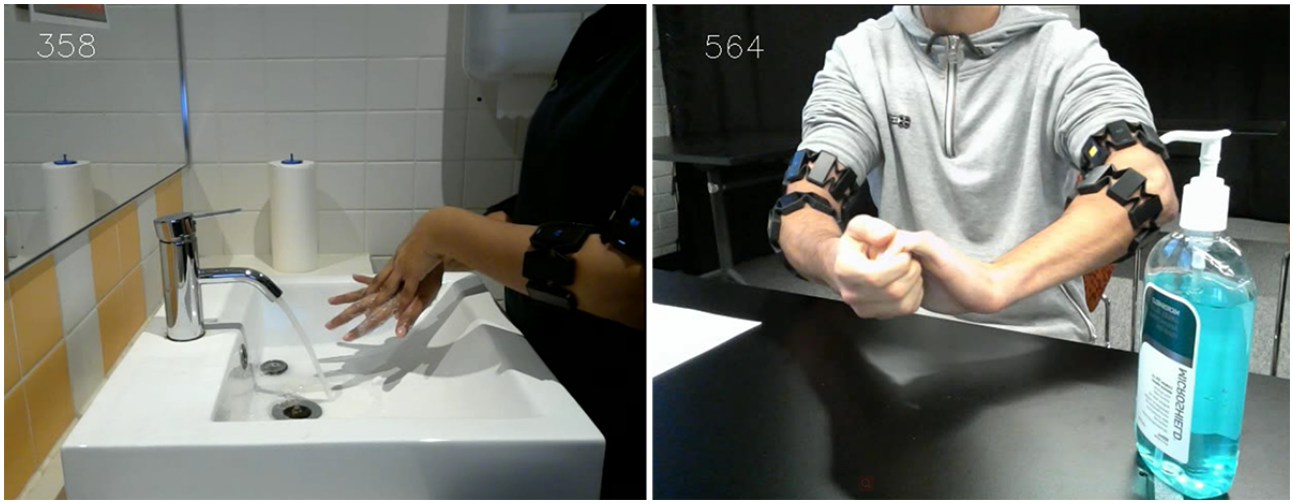
On arrival at our laboratory, we briefed participants on the purpose of the study and obtained their written consent agreeing to participate in our experiment. Then, we asked our participants to wear four Myo armbands on their forearms and arms, as shown in Figure 2. Thus, each participant had an armband in their upper and lower left arm and upper and lower right arm.

Figure 2. Handwash with soap and water (left) and alcohol-based handrub (right).

We subsequently provided training to our participants. We instructed them on how to perform the handwash and handrub procedures. To achieve this, we first explained to them the procedure steps using an instructional poster (Multimedia Appendices 1 and 2) [3]. We then asked participants to watch an instructional video five times while they perform the procedure. We ran this process twice: once for handrub [38] and once for handwash with soap and water [39] in a counterbalanced manner.

After training, we initiated the main experimental task, during which we collected data from the four armbands that participants wore, and we also videotaped participants' hands as a way of capturing the ground truth (Figure 3). Each participant performed 30 sessions: five handwash sessions where they could follow the instructional video, 10 handwash sessions accompanied with the instructional poster, five handrub sessions with the instructional video, and 10 handrub sessions with the instructional poster. We counter balanced the order of the conditions to avoid order effects. Handrub was performed in our laboratory, and handwash was performed at an adjoining lavatory (Figure 2). The overall duration for each participant was approximately 70 min. Each participant was rewarded with an AUD $20 (US $13.4) gift card regardless of their performance.

Figure 3. Screenshots of the video records of handwash with soap and water (left) and alcohol-based handrub (right). The frame number (top left) is used to synchronize video data to the armband data.



Data Analysis

Our first step was to annotate the video recordings manually: we visually inspected all recordings (493 min in total) and annotated each video frame as belonging to one of the 14 handwash steps. All the video recordings were inspected by two data annotators (annotator 1: 16 participants and annotator 2: 4 participants), and annotator 1 checked the annotations made by annotator 2. The interrater reliability was 1.0 for those samples with two annotators. In total, we made 7180 manual annotations, which represented our ground truth. Using these annotations, along with the frame numbers of the video

recordings, we then annotated the armband data with the relevant hand hygiene step label.

Next, we were able to measure the accuracy of the algorithm for classifying the steps of the hand hygiene procedures, which operates as follows: (1) preprocessing, (2) feature extraction, (3) classification, (4) postprocessing, and (5) evaluation.

Preprocessing

Owing to the different frequencies of IMU (50 Hz) and sEMG (200 Hz) data, we downsampled both types of sensor data to 25 Hz. We then normalized the downsampled data by applying z-score normalization and applied a sliding window procedure with a 0.2-second time window with 75% overlap. We also measured the classification accuracy for different sliding window parameters, ranging from 0.1 seconds to 0.4 seconds and from 50% overlapping to 75% overlapping, and found the sliding window with a 0.2-second time window and 75% overlap gave the best classification accuracy.

Feature Extraction

Features extracted from the time domain and the frequency domain are widely used in activity recognition tasks. By summarizing previous work [40-43], we calculated all the selected features for the collected IMU and sEMG data. Table 2 shows the extracted features, and Multimedia Appendix 3 provides the details of features. The dimensionality of the feature vector can vary on different configurations (eg, the number of armbands and sensors). Then, to minimize the required computational power and computation time for the activity recognition task, we used boosting algorithms to select discriminative features [44]. In particular, we used XGBoost [45] to choose the most discriminative subset (top 100) of features according to occurrences of the features in splits; hence, the final feature vector had a dimensionality of 100.

Table 2. Features are extracted from acceleration, gyroscope, orientation, and surface electromyography data.

Features	Study authors
CSD, peak (positive), peak (negative), RMS	McIntosh et al [40]
ACAbsArea, ACAbsCV, ACAbsMean, ACEntropy, ACIQR, ACKur, ACQ1, ACQ3, ACRange, ACSkew, ACVar, DCArea, DCMean, DCPostureDist, DCTotalMean	Munguia Tapia et al [41]
|AL|, |ΔAL|, |ΔAR|, |ΔAR|, |ΔMAV|, AJ, AL, AR, RAJ, RMAV, SAJ, SDAL, SDAR, SRAJ	Xie et al [42]
meanPKT, meanPSD, medainS, medianPKT, medianPSD, stdPKT, stdPSD, stdS	Zhang et al [43]

Classification

We fed the generated feature vectors and the ground truth labels to XGBoost to classify the steps of the WHO hand hygiene procedures shown in Table 1 because XGBoost is also widely used in human activity recognition tasks [46,47]. Other techniques were considered and tested during the pilot study, including random forest and SVM, but they performed worse than XGBoost as same as previous studies [46,47]. XGBoost has been shown to have several other advantages, including high efficiency, low computational cost, supporting parallelization, and robustness to overfitting [47]. For these reasons, we focused on applying XGBoost to recognize the steps of hand hygiene procedures.

Postprocessing

A final step of the data analysis pipeline in our study was smoothing the stream of predictions to remove classification errors because switching from one gesture to another gesture several times per second was not realistic. Instead of HMM used in previous studies, we smoothed the prediction stream through a combination of E.Divisive [33] and majority vote. E.Divisive is a nonparametric multiple change point analysis approach based on hierarchical clustering, which can detect distributional changes from a sequence of data [33]. By fitting the prediction stream into E.Divisive using the ecp library [48], we could estimate the location of the change points and use the change points to segment the prediction stream. Finally, the class of each segment was determined by a majority vote over the predictions in the segment.

Evaluation

We measured the accuracy of both a user-independent model (one model to classify data for all participants) and user-dependent models (one model tuned to each participant). To measure the accuracy of models, we used leave-one-session-out (LOSO) cross-validation for the user-dependent model and leave-one-participant-out (LOPO) cross-validation for the user-independent models, as suggested in the literature [30].

In LOSO cross-validation, we considered 29 hand hygiene sessions from one participant for training and tested the model on the holdout session. We performed this protocol 30 times per participant (thus, each session became the holdout session once) and calculated the average accuracy per participant. We then repeated this procedure for each participant independently.

To evaluate the user-independent model, we used 19 participants' hand hygiene data to train the model and tested the model on the remaining participant. We repeated the cross-validation for every participant in our database and then averaged the results across the 20 runs.

Results

Throughout our results, we make an explicit distinction between the user-independent and the user-dependent models. This is because of the practical implications of choosing one approach over the other. A user-independent model *works for all users* and can be used by an HCW without prior training. A user-dependent model needs to be trained and tuned to each HCW individually. The benefit of the former is that it does not

require further training (also known as bootstrapping), whereas the latter attains improved performance.

Participant Performance

We first considered the participants' performance in terms of hand hygiene quality, and, specifically, we were interested in measuring any variations to their performance. As we gave identical and considerable training to each participant, we expect that their hand hygiene quality is high and consistent. By analyzing our annotations of the recorded videos, we can achieve the following:

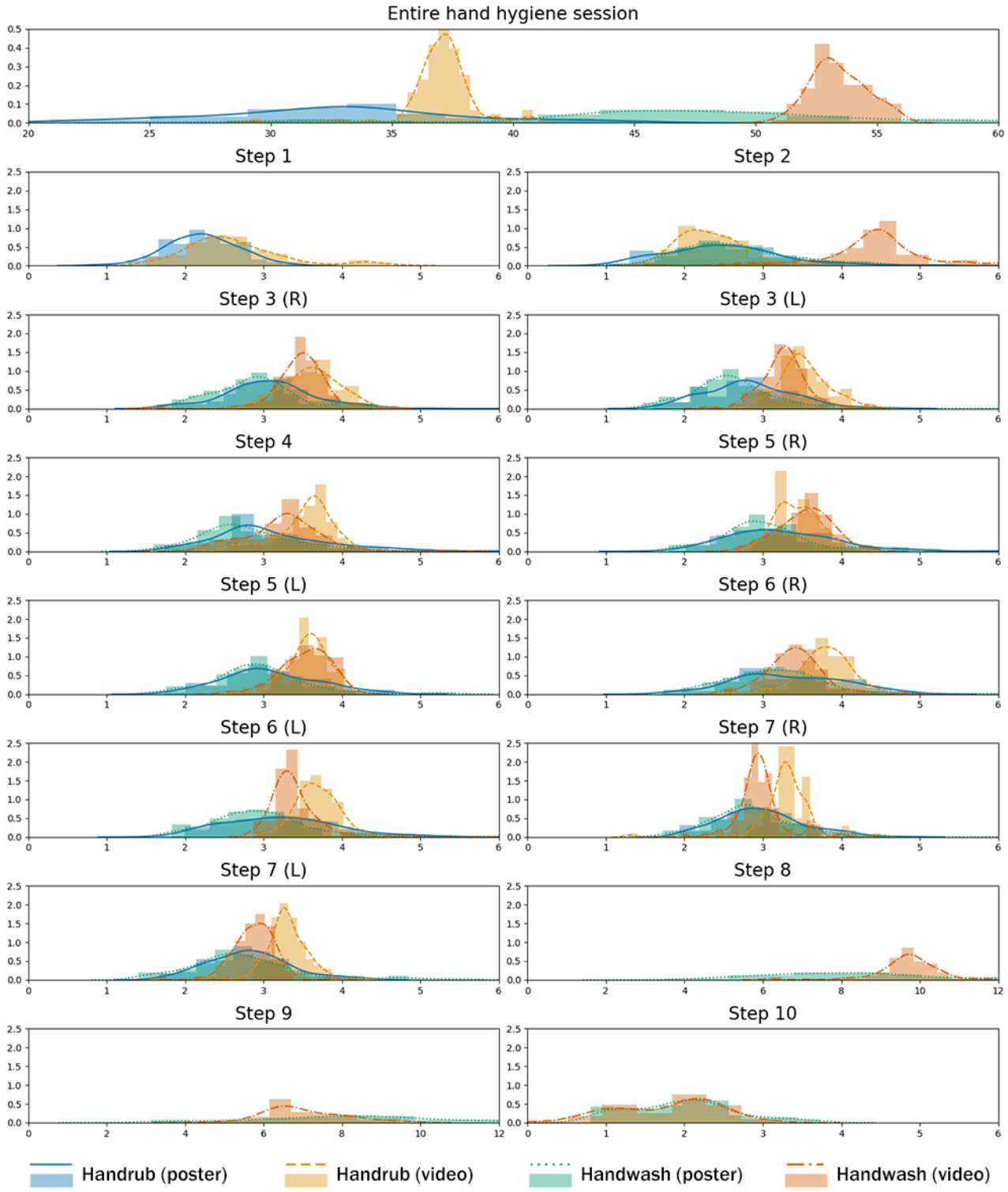
- Measure the exact duration of hand hygiene sessions and steps in those sessions and how they vary across conditions;
- Identify the order in which participants actually performed hand hygiene and whether they missed or skipped steps or performed steps incorrectly; and
- Investigate the presence of possible learning or fatigue effects in the study.

Time to Complete Hand Hygiene

By summarizing the duration of hand hygiene sessions in the top of Figure 4, we observe that duration varies considerably across different conditions as follows: handwash with the instructional video (mean 53.4 seconds, SD 1.1) or poster (mean 47.9 seconds, SD 8.3) and handrub with the video (mean 37.1 seconds, SD 1.4) or poster (mean 32.6 seconds, SD 5.3). We observed this effect for overall timings and the duration of each specific step. A one-way analysis of variance (ANOVA) showed a statistically significant effect of the different conditions on the duration of hand hygiene sessions ($F_{3,596}$=402.83; P<.001). The post hoc Tukey's honestly significant difference test indicated that the duration of the video sessions is longer than that of the poster sessions, and the duration of the handwash sessions is longer than that of the handrub sessions.

Furthermore, we observed that participants' time hygiene varied considerably more (higher standard deviation) in the poster conditions as opposed to the corresponding video conditions. This is not surprising, as the video imposes a certain pace on participants, whereas in the poster condition, participants are free to set their own pace.

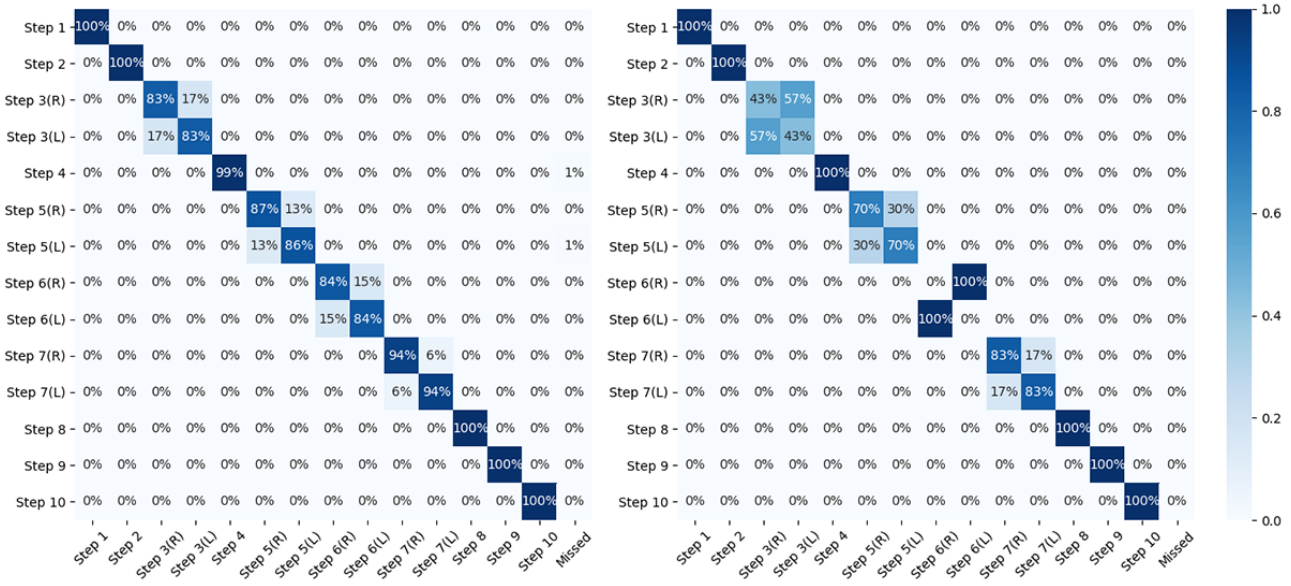
Figure 4. Duration of hand hygiene sessions and inner steps. x-axis: duration (seconds); y-axis: density. L: left; R: right.



Step Sequence Accuracy

By observing the recorded hand hygiene sessions, we noticed that participants did deviate from the defined procedure and had an overall accuracy of 91% (SD 0.16) in terms of complying with the protocol. As shown in Figure 5, we calculated the average participant accuracy as a measure of their compliance for each of the protocol steps. Participants tended to swap those steps that need to be repeated symmetrically with different leading hands, including step 3(R)/3(L), step 5(R)/5(L), step 6(R)/6(L), and step 7(R)/7(L). We also observed that on some occasions, participants missed a step (denoted as rightmost column *Missed* in Figure 5).

Figure 5. Confusion matrix showing the order of steps in hand hygiene sessions: all participants (left) and participant 8 (right). x-axis: performed step; y-axis: excepted step. L: left; R: right.

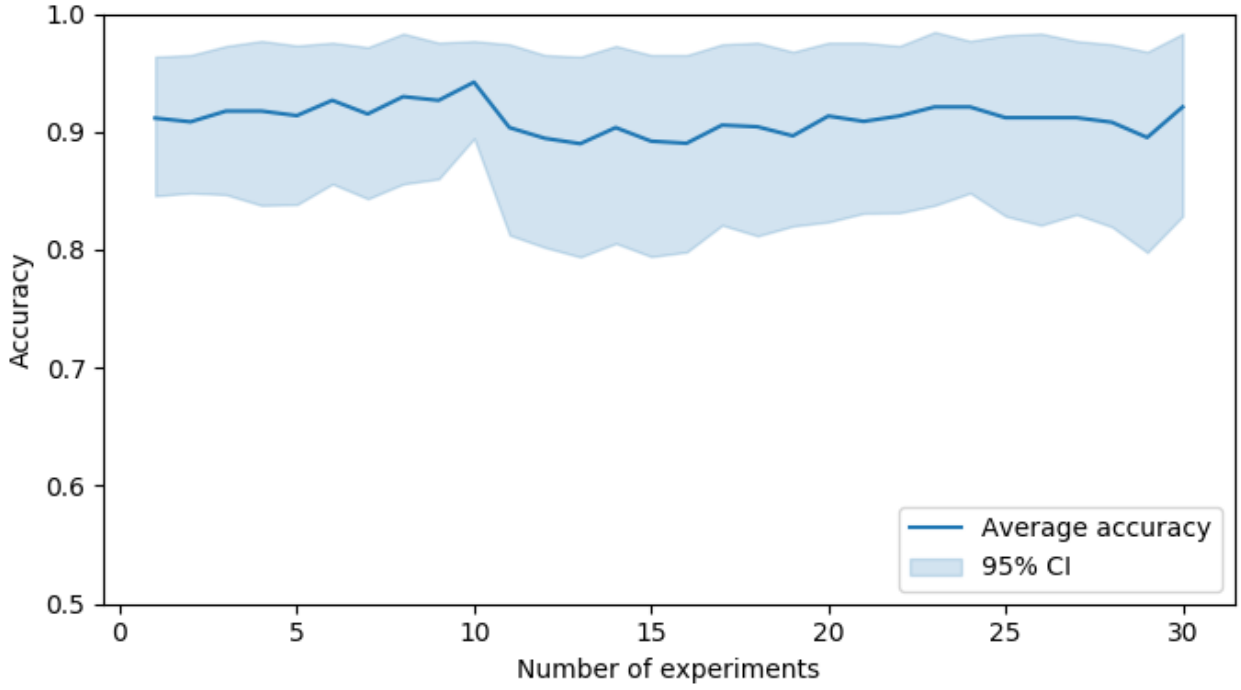


On the right of Figure 5, we show exemplary data collected from participant 8, who swapped the order of steps 6(R) and 6(L). We noted that this participant was right-handed. Overall, we observe that participants are able to follow the standard WHO hand hygiene procedures to a great extent, although participants may accidentally miss some steps or swap some symmetrical steps. We highlight that although these errors were made by participants, they do not affect the accuracy of our classifier. This is because we manually inspected and labeled our video recordings, and therefore, any missed or swapped steps are assigned the correct ground truth label.

Learning and Fatigue Effects

To quantify the existence of any potential learning or fatigue effects, we calculated participants' average accuracy after having completed one procedure, two, three, and so on until 30. The results are shown in Figure 6. The average accuracy remained at around 90% across the range of 1 to 30 completed sessions. A one-way ANOVA showed no statistically significant effect of the number of experiments on the average accuracy ($F_{29,570}$=0.10; P>.99). This indicates that participants did not have a significant increase or decline in performance, which would be suggestive of learning or fatigue, respectively.

Figure 6. Average participant accuracy, according to the number of completed hand hygiene sessions.



Armband Performance

Placement of Armbands

To investigate the performance of the armband, we calculated the accuracy for eight distinct combinations of armband positions: four combinations of two armbands (left arm+right arm, left forearm+right forearm, right arm+right forearm, and left arm+left forearm) and four positions of one armband (right arm, left arm, right forearm, and left forearm). We separately calculated the average accuracy for the classification results (XGBoost) and the smoothed classification results (XGBoost+E.Divisive). At this point, the results are derived using a fusion of all available sensor data (acceleration, gyroscope, orientation, and sEMG) and are shown in Figure 7 (user-dependent models) and Figure 8 (user-independent model). As noted earlier, we used LOSO cross-validation for the user-dependent models and LOPO cross-validation for the user-independent model.

For the user-dependent model, Figure 7 shows the best performance was achieved when using the data from both left forearm and right forearm (XGBoost: mean 95.9%, SD 0.01; and XGBoost+E.Divisive: mean 96.9%, SD 0.01). For situations where only one armband is desirable or available, the best model uses data from the right forearm (XGBoost: mean 93.0%, SD 0.02; and XGBoost+E.Divisive: mean 96.1%, SD 0.01).

Similar result trends were observed in the user-independent model as well, although the user-independent model lags in performance as expected. In Figure 8, the model using the data from both the left forearm and right forearm (XGBoost: mean 85.5%, SD 0.05; and XGBoost+E.Divisive: mean 90.9%, SD 0.04) outperformed other placement combinations. When considering only one armband, the highest classification accuracy was also achieved by the model using the data from the right forearm (XGBoost: mean 72.6%, SD 0.07; and XGBoost+E.Divisive: mean 82.4%, SD 0.08).

Although we observed that the overall classification accuracy of the user-independent model is lower than that of the user-dependent models, we found that the user-independent model using the data from two armbands is able to achieve more than 80% accuracy.

There were several similarities between the results for the user-dependent models and the user-independent model, lending higher robustness to our results. We observed that the classification accuracy increases after smoothing in all cases, shown in Figures 7 and 8 with a two-way ANOVA showing a significant improvement with smoothing (user-dependent model: $F_{1,311}$=197.81; P<.001; and user-independent model: $F_{1,311}$=111.40; P<.001). Figure 9 illustrates how the E.Divisive smoothing algorithm improves the data quality by reducing classification errors. Figure 9 also demonstrates how the classifier can detect steps that are out of order (steps 6(R) and 6(L) were performed in reverse order by the participant), deal with missed steps (step 5(L) was skipped), and detect steps with varying duration. Furthermore, Multimedia Appendices 4 and 5 show the confusion matrices of recognition rates from user-dependent models and user-independent models, respectively.

Figure 7. Classification accuracy of different combinations of armband positions: user-dependent models.

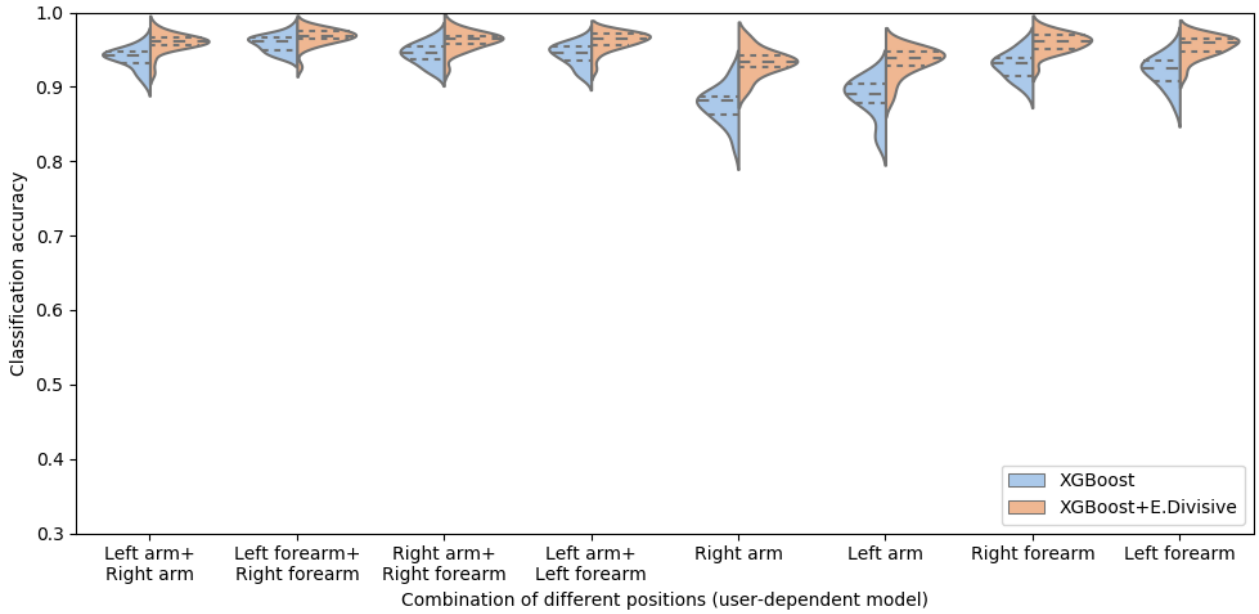


Figure 8. Classification accuracy of different combinations of armband positions: user-independent model.

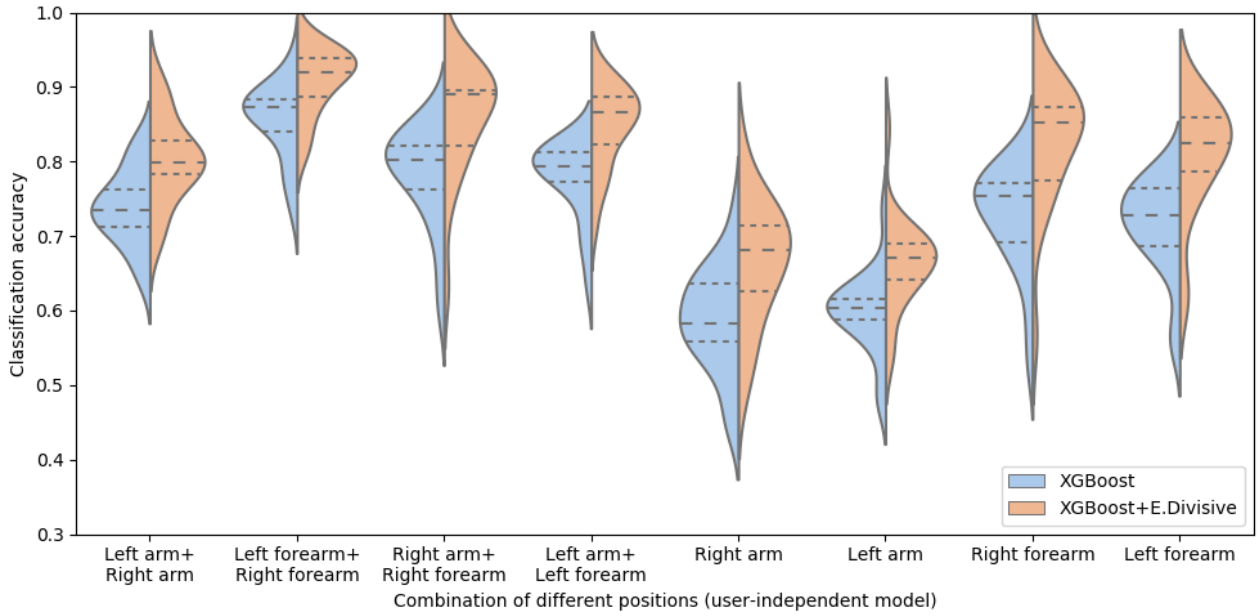
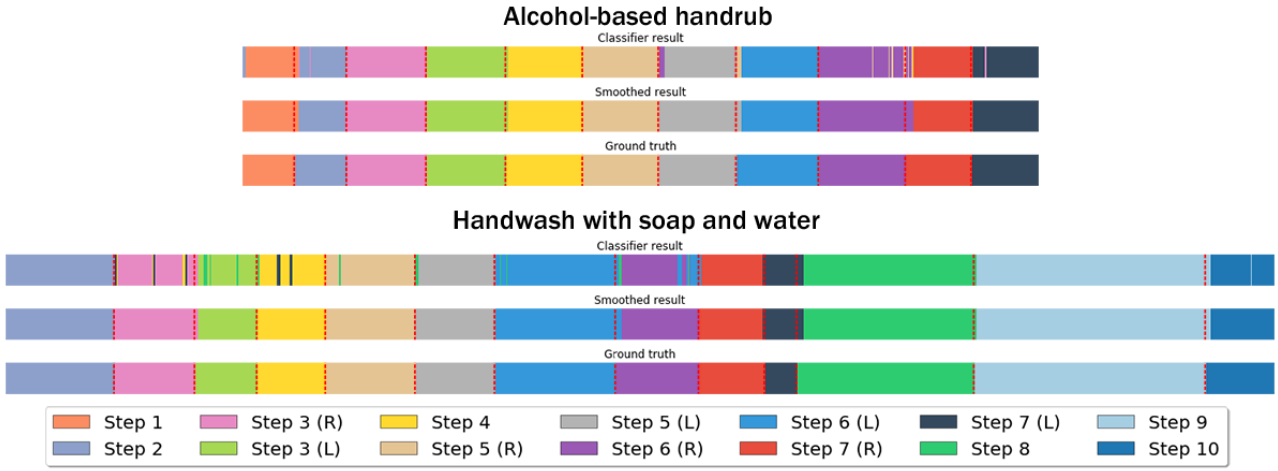


Figure 9. Visualization of how classification and smoothing work for handrub (top) and handwash (bottom). The steps are in accordance with Table 1.

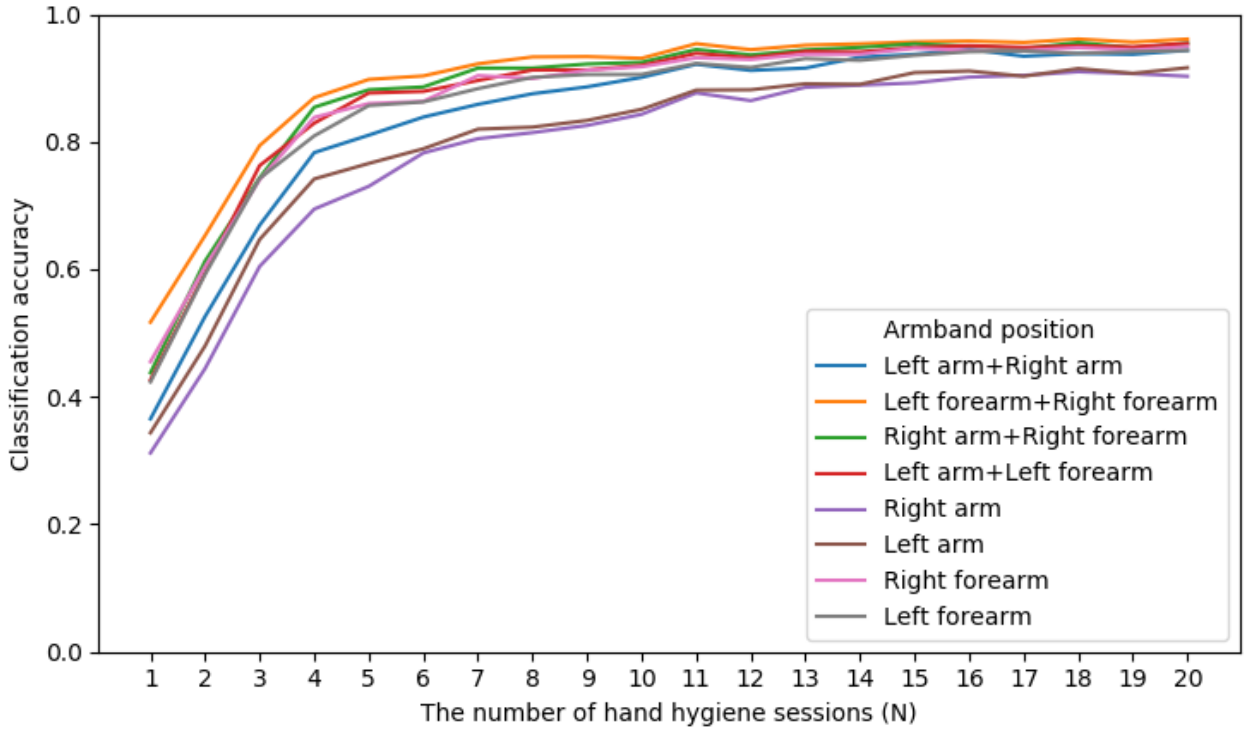


Bootstrapping User-Dependent Models

The results have shown that in scenarios where only one armband is placed on the arm, the performance of the user-independent model drops to below 70% (Figure 8). Considering the low accuracy associated with user-independent models, it would be desirable to use a user-dependent model that can achieve higher classification accuracy of about 95% (Figure 7). However, this requires training data from the individual user, and therefore, the model cannot perform well immediately [49]. This can pose an additional burden on HCWs who need to train the model before it performs well. For this reason, we quantified the amount of training data that is required to achieve a reasonable classification accuracy.

We did this by randomizing the order of both 15 handrub and 15 handwash sessions for each participant, pick 1, 2, ..., *N* hand hygiene sessions as the training set, and test the model on the remaining hand hygiene sessions for that participant. We repeated this process for every participant and calculated the average classification accuracy for each *N*. The results are shown in Figure 10 and illustrate that classification accuracy rapidly increases as the number of training sessions grows from 1 to 6. With six hand hygiene sessions as the training set, the models with one armband placed on participants' arm achieve around 80% accuracy. After 16 hand hygiene sessions, the accuracy of the models solely based on the arm is higher than 90%.

Figure 10. Classification accuracy of user-dependent models using N hand hygiene sessions as the training set.



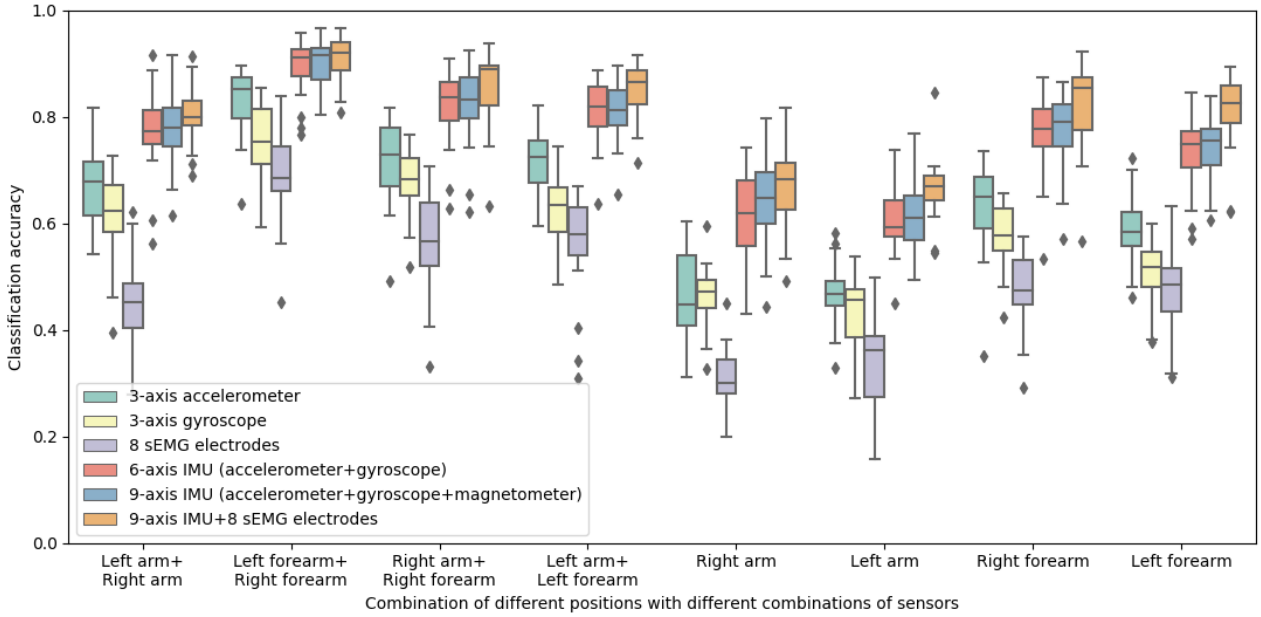
Sensor Importance

It is crucial to minimize the interruptions of HCWs when using wearable technology such as the armbands in our study, for example, to recharge the armband. Therefore, one approach is to minimize the number of sensors on the armband and therefore increase the battery lifetime but reduce the classification accuracy. For this reason, we quantified the trade-off between the power consumption and classification accuracy of the wearable sensors. In addition, one of the major concerns of electronic hand hygiene–monitoring systems is the cost associated with the wearable sensors and corresponding receivers, and therefore, reducing the sensors can reduce the associated costs. Furthermore, another concern is HCWs' acceptance and compliance with hospital regulations (eg, not wearing wristbands or watches), and thus, we expected to minimize the number of sensors and the size of sensor armbands to reduce their impediments and increase HCWs' acceptance.

To quantify this energy vs performance trade-off, we fed the user-independent model with the data from different combinations of sensors in line with the widely used wearable devices. Specifically, we examined the performance of the user-independent model with six different combinations of sensors through LOPO cross-validation. The six combinations are as follows: three-axis accelerometer, three-axis gyroscope, eight sEMG electrodes, six-axis IMU (accelerometer and gyroscope), nine-axis IMU (accelerometer, gyroscope, and magnetometer), and nine-axis IMU + eight sEMG electrodes. Effectively, these are all subsets of the full dataset we collected, and that is why this analysis is possible. The classification accuracy when using different sensor combinations, with different sensor placement, is shown in Figure 11.

Figure 11 shows that, in general, performance suffers when the available data are only a three-axis accelerometer, or a three-axis gyroscope, or eight sEMG electrodes. This finding holds across all armband placements for both one- and two-armband scenarios. The results also showed that if only one sensor is available, then a three-axis accelerometer would be preferable, as it outperforms the others. However, we also observe in Figure 11 a substantial performance gain when a six-axis accelerometer is used—effectively when an accelerometer and gyroscope are combined. Further adding a magnetometer (ie, nine-axis IMU), and subsequently adding the sEMG data, provides mostly marginal performance gains.

Figure 11. Classification accuracy of the user-independent model for different combinations of sensors and different placement. IMU: inertial measurement unit; sEMG: surface electromyography.



Discussion

Flexible Detection

To mimic the real-life scenarios in health care settings, we did not place restrictions on the completion time in our experiments. As a result, we observed that the duration of hand hygiene performance varied considerably when the sessions were accompanied by the instructional poster. Interestingly, performing hand hygiene with posters is common practice in hospitals, and Multimedia Appendices 1 and 2 have been widely used as an approach to promote hand hygiene [50]. In addition, the suggested duration of handwash and handrub from the WHO also varies from 20 to 30 seconds and from 40 to 60 seconds, respectively [3]. Hence, to detect the steps of hand hygiene procedures in realistic scenarios, the system should be adaptive to the varying duration.

Our study also shows that participants did deviate from the defined procedure, especially for the steps that need to be repeated symmetrically with different leading hands. Meanwhile, some participants accidentally missed some of the steps, which is more prevalent in health care settings. For example, Tschudin-Sutter et al [23] reported a compliance rate of only 8.5% for completing all steps of the WHO hand hygiene procedures. Furthermore, the researchers do not have a consensus on the correct order of the hand hygiene procedure steps [51]. Pires et al [51] suggested that HCWs should rub the fingertips first to reduce the probability of contamination. Hence, to monitor the quality of hand hygiene, the system should provide the flexibility of detecting the individual steps within the hand hygiene procedures.

Previous studies adopted HMM to classify the steps of the hand hygiene procedure from feature vectors or smooth classification results [29,30], which assumes that HCWs will perform hand hygiene within a specific time frame and with a predefined sequence. Instead, we smoothed the prediction results through a combination of E.Divisive and majority vote. The results indicated that the classification accuracy increases after smoothing (user-dependent model: $F_{1,311}$=197.81; P<.001; and user-independent model: $F_{1,311}$=111.40; P<.001; see Figures 7 and 8). Our approach can deliver flexibility for detecting hand hygiene in real-life scenarios without assuming that the HCWs follow a predefined sequence within a specific time limit.

Placement Recommendations

To investigate the relationship between the performance of our proposed models and the placement of the armbands, we tested the models with eight placement combinations. For a scenario with two armbands, the model using the data from both the left forearm and the right forearm exhibits the highest classification accuracy (user-dependent model: 96.9% and user-independent model: 90.9%). When using the data from both the right arm and the left arm, the performance decreases to 93.8% and 80%, respectively. However, it can be argued that placing the armbands on the forearm is less hygienic that placing them on the arm.

When only one armband is available, the performance of the user-dependent model still reaches more than 93.3% accuracy regardless of the position of the device (eg, on the lower or upper arm). However, for the user-independent models, the performance drops below 70% when the armband is located on the upper arm (80% accuracy for the lower arm). Thus, a user-dependent model is preferred when only using the data from one armband.

Furthermore, to reduce the costs of data collection, we measured the amount of training data that is required to achieve a reasonable classification accuracy. With six hand hygiene sessions as the training set, the models in the one armband scenario placed on participants' arm achieve around 80% accuracy for the user-dependent model. After 16 hand hygiene sessions, the accuracy of the models is higher than 90%.

However, these models need manually annotated personal training data, which might require considerable costs associated with personnel time and resources.

Overall, for the scenarios allowing HCWs to wear armbands on their forearms, the user-independent models can achieve acceptable classification accuracy (one armband: more than 80% and two armbands: 90.9%). When the hygiene protocol enforces restrictions on HCWs' forearms (eg, the armband is not allowed on the forearm), the user-independent model needs data from both the left arm and right arm because the performance of the user-independent models using one armband is lower than 70%. Owing to this, user-dependent models are necessary and should be trained by six annotated hand hygiene sessions at least to achieve reasonable classification accuracy.

Sensor Recommendations

To minimize the interruptions caused by the wearable sensors (eg, recharge the armbands), one can aim to reduce the number of sensors on the armband so as to increase the battery lifetime. Therefore, we quantified the trade-off between the number of wearable sensors and the corresponding classification accuracy. Our results showed that if only one sensor is available, then a three-axis accelerometer would be preferable, as it outperforms the other sensors. After combining an accelerometer and a gyroscope, the model shows a substantial increase. Further adding a magnetometer (ie, nine-axis IMU), and subsequently adding the sEMG data, provides marginal performance gains at the cost of substantial power consumption. Other studies also showed that after adding sEMG data, the performance of user-independent models does not exhibit a considerable increase [52,53].

These results point to two sensor combinations that appear to minimize power consumption and cost while maximizing performance. For a single-sensor configuration, a three-axis accelerometer is preferable. For multisensor configurations, a six-axis IMU (ie, accelerometer+gyroscope) is the parsimonious combination of choice. We also observed that for single-armband scenarios, the classification accuracy with single-sensor data is largely below 60%, and therefore, a multisensor configuration is appropriate where only one armband is available.

Feedback

One of the key steps for medical students and HCWs to acquire the clinical techniques is providing feedback on their performance in given activities [54]. Before receiving feedback regarding their performance of hand hygiene, the first step should be to reveal the gap between the optimal and actual performance.

In practice, the duration of a hand hygiene procedure is considered as the key indicator of quality [22,55]. Meantime, as mentioned by Arias et al [21] and Tschudin-Sutter et al [23], the noncompliance with all steps of hand hygiene procedures results in failure to cover all skin surfaces; hence, an automated monitoring approach should also provide information about the sequence of hand hygiene steps actually followed. Owing to the flexibility of our system, we can detect a hand hygiene procedure with different duration and orders. Therefore, we can provide feedback to medical students and HCWs on the duration of hand hygiene procedure (eg, which steps are performed, how much time they spent on the specific steps, and whether they should prolong the duration of the specific step) and the sequence of hand hygiene procedures (eg, whether there is any missed step and whether performed steps are in the correct order).

Through instructional applications, such feedback can be based on the individual performance of hand hygiene so that the trainees can receive more detailed and personalized instructions during the training period. This approach can be used to investigate the performance of HCWs' daily hand hygiene events. HCWs can further improve their hand hygiene techniques through timely and periodic feedback. Similarly, administrators can also have summarized information to help them quantify hand hygiene quality.

Limitations

In this study, we recruited participants through our university's mailing list and collected data in a controlled laboratory setting and did not conduct a field study in health care settings. However, it was necessary to conduct a laboratory study first to provide a semicontrolled setting for instrument validation. We also provided substantial training to participants, and we did observe their performance to be higher than 90%, suggesting that our training was effective in ensuring they follow the WHO guidelines during the study.

Another limitation is that we did not measure the effect of handedness on classifier performance because of the small number of left-handed participants (1/20, 5% participants). However, as mentioned by Galluzzi et al [34], hand dominance is not crucial to hand hygiene detection because of symmetric arm movements.

In addition, we limited our approach to measuring the performance of hand hygiene quality to Myo armbands, which have officially ended as of October 2018. Nevertheless, we argue that a similar approach should be applicable with other wearable devices that bear IMU + electromyography sensors, so long as our proposed analysis is adopted.

Conclusions

In this paper, we evaluated the feasibility of using sensor armbands (Myo armband) to assess the HCWs' compliance with the WHO hand hygiene guidelines, which are considered as the proxy measures of the quality of hand hygiene. Our results showed the classification performance of 97% average accuracy for the user-dependent model and 91% average accuracy for the user-independent model. In addition, by investigating the performance of individual models with different sizes of training data, we found that training the user-dependent model with six annotated hand hygiene events can provide more than 80% accuracy with the data from one armband placed on the users' upper arms. We also investigated the performance of different sensor combinations and found that the combination of an accelerometer and a gyroscope achieves the balance between the classification performance, power consumption, and cost. Our findings contribute to building mechanisms to quantify the quality of hand hygiene procedures using a sensor armband.

Acknowledgments

This work was partially funded by the Australian Research Council Discovery Project DP190102627 and the National Health and Medical Research Council grant 1170937. CW was supported by a PhD scholarship provided by the Australian Commonwealth Government Research Training Program.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Standard World Health Organization procedure of handwash with soap and water. Source: WHO Guidelines on Hand Hygiene in Health Care: First Global Patient Safety Challenge Clean Care is Safer Care [3]
[PDF File (Adobe PDF File), 457 KB - mhealth_v8i3e17001_app1.pdf]

Multimedia Appendix 2
Standard World Health Organization procedure of alcohol-based handrub. Source: WHO Guidelines on Hand Hygiene in Health Care: First Global Patient Safety Challenge Clean Care is Safer Care [3]
[PDF File (Adobe PDF File), 426 KB - mhealth_v8i3e17001_app2.pdf]

Multimedia Appendix 3
Description of the selected features.
[XLSX File (Microsoft Excel File), 11 KB - mhealth_v8i3e17001_app3.xlsx]

Multimedia Appendix 4
Confusion matrix showing the recognition rates of user-dependent models according to different placement combinations.
[PNG File , 287 KB - mhealth_v8i3e17001_app4.png]

Multimedia Appendix 5
Confusion matrix showing the recognition rates of user-independent models according to different placement combinations.
[PNG File , 404 KB - mhealth_v8i3e17001_app5.png]

References

1. Allegranzi B, Pittet D. Role of hand hygiene in healthcare-associated infection prevention. J Hosp Infect 2009 Dec;73(4):305-315. [doi: 10.1016/j.jhin.2009.04.019] [Medline: 19720430]
2. Loveday H, Wilson J, Pratt R, Golsorkhi M, Tingle A, Bak A, UK Department of Health. epic3: national evidence-based guidelines for preventing healthcare-associated infections in NHS hospitals in England. J Hosp Infect 2014 Jan;86(Suppl 1):S1-70. [doi: 10.1016/S0195-6701(13)60012-2] [Medline: 24330862]
3. World Health Organization. WHO Guidelines on Hand Hygiene in Health Care: First Global Patient Safety Challenge Clean Care is Safer Care. Geneva, Switzerland: World Health Organization; 2009.
4. Klevens RM, Edwards JR, Richards CL, Horan TC, Gaynes RP, Pollock DA, et al. Estimating health care-associated infections and deaths in US hospitals, 2002. Public Health Rep 2007;122(2):160-166 [FREE Full text] [doi: 10.1177/003335490712200205] [Medline: 17357358]
5. Allegranzi B, Nejad SB, Combescure C, Graafmans W, Attar H, Donaldson L, et al. Burden of endemic health-care-associated infection in developing countries: systematic review and meta-analysis. Lancet 2011 Jan 15;377(9761):228-241. [doi: 10.1016/S0140-6736(10)61458-4] [Medline: 21146207]
6. Ling ML, Apisarnthanarak A, Madriaga G. The burden of healthcare-associated infections in Southeast Asia: a systematic literature review and meta-analysis. Clin Infect Dis 2015 Jun 1;60(11):1690-1699. [doi: 10.1093/cid/civ095] [Medline: 25676799]
7. Magill SS, Edwards JR, Bamberg W, Beldavs ZG, Dumyati G, Kainer MA, Emerging Infections Program Healthcare-Associated Infections and Antimicrobial Use Prevalence Survey Team. Multistate point-prevalence survey of health care-associated infections. N Engl J Med 2014 Mar 27;370(13):1198-1208 [FREE Full text] [doi: 10.1056/NEJMoa1306801] [Medline: 24670166]
8. Zarb P, Coignard B, Griskeviciene J, Muller A, Vankerckhoven V, Weist K, National Contact Points for the ECDC pilot point prevalence survey, Hospital Contact Points for the ECDC pilot point prevalence survey. The European Centre for Disease Prevention and Control (ECDC) pilot point prevalence survey of healthcare-associated infections and antimicrobial use. Euro Surveill 2012 Nov 15;17(46):pii: 20316 [FREE Full text] [doi: 10.2807/ese.17.46.20316-en] [Medline: 23171822]

9. De Angelis G, Murthy A, Beyersmann J, Harbarth S. Estimating the impact of healthcare-associated infections on length of stay and costs. Clin Microbiol Infect 2010 Dec;16(12):1729-1735 [FREE Full text] [doi: 10.1111/j.1469-0691.2010.03332.x] [Medline: 20673257]
10. Lamarsalle L, Hunt B, Schauf M, Szwarcensztein K, Valentine W. Evaluating the clinical and economic burden of healthcare-associated infections during hospitalization for surgery in France. Epidemiol Infect 2013 Dec;141(12):2473-2482 [FREE Full text] [doi: 10.1017/S0950268813000253] [Medline: 23445665]
11. Graves N, Halton K, Paterson D, Whitby M. Economic rationale for infection control in Australian hospitals. Healthc Infect 2009;14(3):81-88 [FREE Full text] [doi: 10.1071/hi09010]
12. Pittet D. Improving adherence to hand hygiene practice: a multidisciplinary approach. Emerg Infect Dis 2001;7(2):234-240 [FREE Full text] [doi: 10.3201/eid0702.010217] [Medline: 11294714]
13. Pittet D, Allegranzi B, Boyce J, World Health Organization World Alliance for Patient Safety First Global Patient Safety Challenge Core Group of Experts. The World Health Organization Guidelines on Hand Hygiene in Health Care and their consensus recommendations. Infect Control Hosp Epidemiol 2009 Jul;30(7):611-622. [doi: 10.1086/600379] [Medline: 19508124]
14. Boyce JM. Electronic monitoring in combination with direct observation as a means to significantly improve hand hygiene compliance. Am J Infect Control 2017 May 1;45(5):528-535. [doi: 10.1016/j.ajic.2016.11.029] [Medline: 28456322]
15. Pittet D, Simon A, Hugonnet S, Pessoa-Silva CL, Sauvan V, Perneger TV. Hand hygiene among physicians: performance, beliefs, and perceptions. Ann Intern Med 2004 Jul 6;141(1):1-8. [doi: 10.7326/0003-4819-141-1-200407060-00008] [Medline: 15238364]
16. Erasmus V, Daha TJ, Brug H, Richardus JH, Behrendt MD, Vos MC, et al. Systematic review of studies on compliance with hand hygiene guidelines in hospital care. Infect Control Hosp Epidemiol 2010 Mar;31(3):283-294. [doi: 10.1086/650451] [Medline: 20088678]
17. Ellison RT, Barysauskas CM, Rundensteiner EA, Wang D, Barton B. A prospective controlled trial of an electronic hand hygiene reminder system. Open Forum Infect Dis 2015 Dec;2(4):ofv121 [FREE Full text] [doi: 10.1093/ofid/ofv121] [Medline: 26430698]
18. Levy A, O'Maley JB, Parker FJ. Google Patents. 2002. Handwash Monitoring System URL: https://patents.google.com/patent/US6426701 [accessed 2020-02-09]
19. Staats BR, Dai H, Hofmann DA, Milkman KL. Motivating process compliance through individual electronic monitoring: an empirical examination of hand hygiene in healthcare. Manag Sci 2017;63(5):1563-1585 [FREE Full text] [doi: 10.2139/ssrn.2964542]
20. World Health Organization. How to Handrub? / How to Handwash? URL: https://www.who.int/gpsc/tools/GPSC-HandRub-Wash.pdf [accessed 2020-02-09]
21. Arias AV, Garcell HG, Ochoa YR, Arias KF, Miranda FR. Assessment of hand hygiene techniques using the World Health Organization's six steps. J Infect Public Health 2016;9(3):366-369 [FREE Full text] [doi: 10.1016/j.jiph.2015.11.006] [Medline: 26707705]
22. Korhonen A, Ojanperä H, Puhto T, Järvinen R, Kejonen P, Holopainen A. Adherence to hand hygiene guidelines - significance of measuring fidelity. J Clin Nurs 2015 Nov;24(21-22):3197-3205. [doi: 10.1111/jocn.12969] [Medline: 26331437]
23. Tschudin-Sutter S, Sepulcri D, Dangel M, Schuhmacher H, Widmer AF. Compliance with the World Health Organization hand hygiene technique: a prospective observational study. Infect Control Hosp Epidemiol 2015 Apr;36(4):482-483. [doi: 10.1017/ice.2014.82] [Medline: 25782907]
24. Szilágyi L, Haidegger T, Lehotsky A, Nagy M, Csonka EA, Sun X, et al. A large-scale assessment of hand hygiene quality and the effectiveness of the 'WHO 6-steps'. BMC Infect Dis 2013 May 30;13:249 [FREE Full text] [doi: 10.1186/1471-2334-13-249] [Medline: 23718728]
25. Llorca DF, Parra I, Sotelo M, Lacey G. A vision-based system for automatic hand washing quality assessment. Mach Vis Appl 2011;22:219-234 [FREE Full text] [doi: 10.1007/s00138-009-0234-7]
26. Xia B, Dahyot R, Ruttle J, Caulfield D, Lacey G. Hand Hygiene Poses Recognition with RGB-D Videos. In: Proceedings of the 17th Irish Machine Vision and Image Processing Conference. 2015 Presented at: IMVIP'15; August 26-28, 2015; Dublin, Ireland.
27. Hoey J, Poupart P, Bertoldi A, Craig T, Boutilier C, Mihailidis A. Automated handwashing assistance for persons with dementia using video and a partially observable Markov decision process. Comput Vis Image Underst 2010;114(5):503-519 [FREE Full text] [doi: 10.1016/j.cviu.2009.06.008]
28. Armellino D, Hussain E, Schilling ME, Senicola W, Eichorn A, Dlugacz Y, et al. Using high-technology to enforce low-technology safety measures: the use of third-party remote video auditing and real-time feedback in healthcare. Clin Infect Dis 2012 Jan 1;54(1):1-7 [FREE Full text] [doi: 10.1093/cid/cir773] [Medline: 22109950]
29. Kutafina E, Laukamp D, Bettermann R, Schroeder U, Jonas S. Wearable sensors for eLearning of manual tasks: using forearm EMG in hand hygiene training. Sensors (Basel) 2016 Aug 3;16(8):pii: E1221 [FREE Full text] [doi: 10.3390/s16081221] [Medline: 27527167]
30. Li H, Chawla S, Li R, Jain S, Abowd G, Starner T, et al. Wristwash: Towards Automatic Handwashing Assessment Using a Wrist-worn Device. In: Proceedings of the 2018 ACM International Symposium on Wearable Computers. 2018 Presented

at: ISWC'18; October 8 - 12, 2018; Singapore, Singapore p. 132-139 URL: https://doi.org/10.1145/3267242.3267247 [doi: 10.1145/3267242.3267247]
31. Mondol MA, Stankovic JA. Harmony: A Hand Wash Monitoring and Reminder System using Smart Watches. In: Proceedings of the 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services on 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services. New York, NY, USA: Association for Computing Machinery; 2015 Presented at: MOBIQUITOUS'15; July 22-24, 2015; Coimbra, Portugal p. 11-20 URL: https://doi.org/10.4108/eai.22-7-2015.2260042 [doi: 10.4108/eai.22-7-2015.2260042]
32. Shhedi ZA, Moldoveanu A, Moldoveanu F, Taslitchi C. Real-time Hand Hygiene Monitoring System for HAI Prevention. In: Proceedings of the 2015 E-Health and Bioengineering Conference. 2015 Presented at: EHB'15; November 19-21, 2015; Iasi, Romania URL: https://doi.org/10.1109/EHB.2015.7391474 [doi: 10.1109/ehb.2015.7391474]
33. Matteson DS, James NA. A nonparametric approach for multiple change point analysis of multivariate data. J Am Stat Assoc 2014;109(505):334-345 [FREE Full text] [doi: 10.1080/01621459.2013.849605]
34. Galluzzi V, Herman T, Polgreen P. Hand Hygiene Duration and Technique Recognition Using Wrist-Worn Sensors. In: Proceedings of the 14th International Conference on Information Processing in Sensor Networks. 2015 Presented at: IPSN'15; April 13 - 16, 2015; Seattle, Washington p. 106-117 URL: https://doi.org/10.1145/2737095.2737106 [doi: 10.1145/2737095.2737106]
35. Myo Support. Myo Connect, SDK and Firmware Downloads URL: https://support.getmyo.com/hc/en-us/articles/360018409792-Myo-Connect-SDK-and-firmware-downloads [accessed 2022-02-09]
36. Rosenstein N. myo-python - Read the Docs. myo-python - Python bindings for the Myo SDK URL: https://myo-python.readthedocs.io/en/latest/ [accessed 2020-02-09]
37. Wang C, Sarsenbayeva Z, Luo C, Goncalves J, Kostakos V. Improving Wearable Sensor Data Quality Using Context Markers. In: Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers. 2019 Presented at: UbiComp/ISWC'19 Adjunct; September 9 - 13, 2019; London, United Kingdom p. 598-601 URL: https://doi.org/10.1145/3341162.3349334 [doi: 10.1145/3341162.3349334]
38. World Health Organization. YouTube. 2015 Oct 20. WHO: How to Handrub? With Alcohol-based Formulation URL: https://www.youtube.com/watch?v=ZnSjFr6J9HI [accessed 2020-02-09]
39. World Health Organization. YouTube. 2015 Oct 20. WHO: How to Handwash? With Soap and Water URL: https://www.youtube.com/watch?v=3PmVJQUCm4E [accessed 2020-02-09]
40. McIntosh J, McNeill C, Fraser M, Kerber F, Markus L, Antonio K. EMPress: Practical Hand Gesture Classification with Wrist-Mounted EMG and Pressure Sensing. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: Association for Computing Machinery; 2016 Presented at: CHI'16; May 7 - 12, 2016; San Jose, California, USA p. 2332-2342 URL: https://doi.org/10.1145/2858036.2858093 [doi: 10.1145/2858036.2858093]
41. Tapia EM. Alumni — MIT Media. 2008 Jun. Using Machine Learning for Real-time Activity Recognition and Estimation of Energy Expenditure URL: http://alumni.media.mit.edu/~emunguia/pdf/PhDThesisMunguiaTapia08.pdf [accessed 2020-02-09]
42. Xie R, Cao J. Accelerometer-based hand gesture recognition by neural network and similarity matching. IEEE Sensors J 2016;16(11):4537-4545 [FREE Full text] [doi: 10.1109/jsen.2016.2546942]
43. Zhang Y, Gu T, Luo C, Kostakos V, Seneviratne A. FinDroidHR. Proc ACM Interact Mob Wearable Ubiquitous Technol 2018;2(1):1-42 [FREE Full text] [doi: 10.1145/3191788]
44. Su X, Tong H, Ji P. Activity recognition with smartphone sensors. Tinshhua Sci Technol 2014;19(3):235-249 [FREE Full text] [doi: 10.1109/tst.2014.6838194]
45. Chen T, Carlos G. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: Association for Computing Machinery; 2016 Presented at: KDD'16; August 13 - 17, 2016; San Francisco, California, USA p. 785-794 URL: https://doi.org/10 [doi: 10.1145/2939672.2939785]
46. Wang L, Gjoreski H, Murao K, Okita T, Roggen D. Summary of the Sussex-Huawei Locomotion-Transportation Recognition Challenge. In: Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. New York, NY, USA: Association for Computing Machinery; 2018 Presented at: UbiComp'18; October 8 - 12, 2018; Singapore, Singapore p. 1521-1530 URL: https://doi.org/10.1145/3267305 [doi: 10.1145/3267305.3267519]
47. Zhang W, Zhao X, Li Z. A comprehensive study of smartphone-based indoor activity recognition via Xgboost. IEEE Access 2019;7:80027-80042 [FREE Full text] [doi: 10.1109/access.2019.2922974]
48. James NA, Matteson DS. ecp: An R Package for Nonparametric Multiple Change Point Analysis of Multivariate Data. J Stat Soft 2015;62(7):1-29 [FREE Full text] [doi: 10.18637/jss.v062.i07]
49. Bishop CM. Pattern Recognition and Machine Learning. New York: Springer; 2006.
50. Pittet D. Commentary: hand hygiene: it's all about when and how. Infect Control Hosp Epidemiol 2008 Oct;29(10):957-959. [doi: 10.1086/592218] [Medline: 18754740]

51. Pires D, Bellissimo-Rodrigues F, Soule H, Gayet-Ageron A, Pittet D. Revisiting the WHO 'How to Handrub' hand hygiene technique: fingertips first? Infect Control Hosp Epidemiol 2017 Feb;38(2):230-233. [doi: 10.1017/ice.2016.241] [Medline: 27817763]
52. Georgi M, Christoph A, Tanja S. Recognizing Hand and Finger Gestures with IMU based Motion and EMG based Muscle Activity Sensing. In: Proceedings of the International Conference on Bio-inspired Systems and Signal Processing. 2015 Presented at: BIOSTEC'2015; January 12-15, 2015; Lisbon, Portugal p. 99-108 URL: https://doi.org/10.5220/0005276900990108 [doi: 10.5220/0005276900990108]
53. Jiang S, Lv B, Guo W, Zhang C, Wang H, Sheng X, et al. Feasibility of wrist-worn, real-time hand, and surface gesture recognition via sEMG and IMU Sensing. IEEE Trans Ind Inf 2018;14(8):3376-3385 [FREE Full text] [doi: 10.1109/tii.2017.2779814]
54. Ende J. Feedback in clinical medical education. J Am Med Assoc 1983 Aug 12;250(6):777-781. [doi: 10.1001/jama.1983.03340060055026] [Medline: 6876333]
55. Goroncy-Bermes P, Koburger T, Meyer B. Impact of the amount of hand rub applied in hygienic hand disinfection on the reduction of microbial counts on hands. J Hosp Infect 2010 Mar;74(3):212-218. [doi: 10.1016/j.jhin.2009.09.018] [Medline: 20061058]

Abbreviations

ANOVA: analysis of variance
HAI: health care–associated infection
HCWs: health care workers
HMM: Hidden Markov Model
IMU: inertial measurement unit
LOPO: leave-one-participant-out
LOSO: leave-one-session-out
sEMG: surface electromyography
SVM: support vector machine
WHO: World Health Organization
XGBoost: eXtreme Gradient Boosting

Edited by G Eysenbach; submitted 11.11.19; peer-reviewed by E Kutafina, H Li; comments to author 02.12.19; revised version received 20.12.19; accepted 24.01.20; published 26.03.20.

Please cite as:
Wang C, Sarsenbayeva Z, Chen X, Dingler T, Goncalves J, Kostakos V
Accurate Measurement of Handwash Quality Using Sensor Armbands: Instrument Validation Study
JMIR Mhealth Uhealth 2020;8(3):e17001
URL: http://mhealth.jmir.org/2020/3/e17001/
doi:10.2196/17001
PMID:32213469

Original Paper

Patterns of Sedentary Time and Quality of Life in Women With Fibromyalgia: Cross-Sectional Study From the al-Ándalus Project

Blanca Gavilán-Carrera[1,2,3], BSc; Víctor Segura-Jiménez[4,5], PhD; Pedro Acosta-Manzano[1,2,3], BSc; Milkana Borges-Cosic[1,2,3], PhD; Inmaculada C Álvarez-Gallardo[1,3,4], PhD; Manuel Delgado-Fernández[1,2,3], PhD

[1]Department of Physical Education and Sport, Faculty of Sport Sciences, University of Granada, Granada, Spain

[2]Sport and Health University Research Institute (iMUDS), Granada, Spain

[3]Physical Activity for Health Promotion Research Group (PA-HELP), Granada, Spain

[4]Department of Physical Education, Faculty of Education Sciences, University of Cádiz, Puerto Real, Cádiz, Spain

[5]Biomedical Research and Innovation Institute of Cádiz (INiBICA) Research Unit, Puerta del Mar University Hospital, University of Cádiz, Cádiz, Spain

Corresponding Author:
Víctor Segura-Jiménez, PhD
Department of Physical Education
Faculty of Education Sciences
University of Cádiz
Av República Saharaui s/n
11519
Puerto Real, Cádiz
Spain
Phone: 34 956016219
Email: victor.segura@uca.es

Abstract

Background: Sedentary time (ST) has been associated with detrimental health outcomes in fibromyalgia. Previous evidence in the general population has shown that not only is the total amount of ST harmful but the pattern of accumulation of sedentary behaviors is also relevant to health, with prolonged unbroken periods (ie, bouts) being particularly harmful.

Objective: To examine the association of the patterns of ST with health-related quality of life (HRQoL) in women with fibromyalgia and to test whether these associations are independent of moderate-to-vigorous physical activity (MVPA).

Methods: A total of 407 women (mean 51.4 years of age [SD 7.6]) with fibromyalgia participated. ST and MVPA were measured with triaxial accelerometry. The percentage of ST accumulated in bouts and the frequency of sedentary bouts of different lengths (≥10 min, ≥20 min, ≥30 min, and ≥60 min) were obtained. Four groups combining total ST and sedentary bout duration (≥30 min) were created. We assessed HRQoL using the 36-item Short-Form Health Survey (SF-36).

Results: A greater percentage of ST spent in all bout lengths was associated with worsened physical function, bodily pain, vitality, social function, and physical component summary (PCS) (all $P<.05$). In addition, a higher percentage of ST in bouts of 60 minutes or more was related to worsened physical role ($P=.04$). A higher frequency of bouts was negatively associated with physical function, social function, the PCS (≥30 min and ≥60 min), physical role (≥60 min), bodily pain (≥60 min), and vitality (≥20 min, ≥30 min, and ≥60 min) (all $P<.05$). Overall, for different domains of HRQoL, these associations were independent of MVPA for higher bout lengths. Patients with high total ST and high sedentary bout duration had significantly worsened physical function (mean difference 8.73 units, 95% CI 2.31-15.15; independent of MVPA), social function (mean difference 10.51 units, 95% CI 2.59-18.44; not independent of MVPA), and PCS (mean difference 2.71 units, 95% CI 0.36-5.06; not independent of MVPA) than those with low ST and low sedentary bout duration.

Conclusions: Greater ST in prolonged periods of any length and a higher frequency of ST bouts, especially in longer bout durations, are associated with worsened HRQoL in women with fibromyalgia. These associations were generally independent of MVPA.

(JMIR Mhealth Uhealth 2020;8(3):e14538) doi:10.2196/14538

KEYWORDS

GT3X+; accelerometry; sedentary behavior; symptomatology

Introduction

Fibromyalgia is a chronic and heterogeneous condition characterized by pain as the dominant symptom, which is frequently accompanied by fatigue, sleep disorders, or cognitive impairment [1]. Fibromyalgia patients, who tend to be highly sedentary, usually reduce their physical activity (PA) levels in order to avoid an aggravation of their symptomatology [2,3]. However, adopting this behavior might trigger a worsening of their condition [4-8]. Importantly, the risks of a sedentary lifestyle are present irrespective of the PA performed [9,10]. Considering that few patients with fibromyalgia fulfil the recommended level of moderate-to-vigorous PA (MVPA) [11], these patients are at an increased health risk not only for being highly sedentary but also for being inactive [12,13]. In the management of fibromyalgia, a graduated approach first focused on nonpharmacological modalities, and the improvement of health-related quality of life (HRQoL) is currently recommended [14]. Therefore, greater insights on how modifiable factors, such as daily sedentary time (ST) and PA, are related to HRQoL among these patients are warranted.

Emerging evidence in the general population has demonstrated that not only the total amount of ST but also the pattern of accumulation of sedentary behaviors is relevant to health [15-17]. Prolonged, unbroken periods (ie, bouts) of ST might be particularly harmful [15,16] due to its relationship with detrimental effects on the metabolism [15-17]. In fibromyalgia, Ellingson et al demonstrated that both total ST, but especially sustained ST, can negatively influence pain modulation processes [18]. In addition, the frequency of sedentary bouts seems to be linked to health outcomes, with frequent interruptions in prolonged ST (ie, breaks) being beneficially related to markers of metabolic risk [19]. Although current PA recommendations emphasize the importance of reducing total ST [20], there is no information on how sedentary behavior patterns (ie, bout duration and frequency) should be modified to maximize health benefits. Sedentary patterns have been typically collected by accelerometers in the research field [17]. In contrast, mobile health (ie, mHealth) tools are more user-friendly devices that are widely used by consumers to track daily activity. These wearable devices, however, do not usually offer sedentary behavior information to users, although the inclusion of accelerometer sensors in wearable devices [21] would make it possible. Therefore, the analysis of the impact of different patterns of sustained ST on HRQoL in fibromyalgia could help in (1) the development of recommendations to reduce overall ST and in the interruption of potentially harmful bout lengths and (2) the implementation of future mHealth tools that deliver actual sedentary pattern information and potentially encourage this population to break prolonged ST.

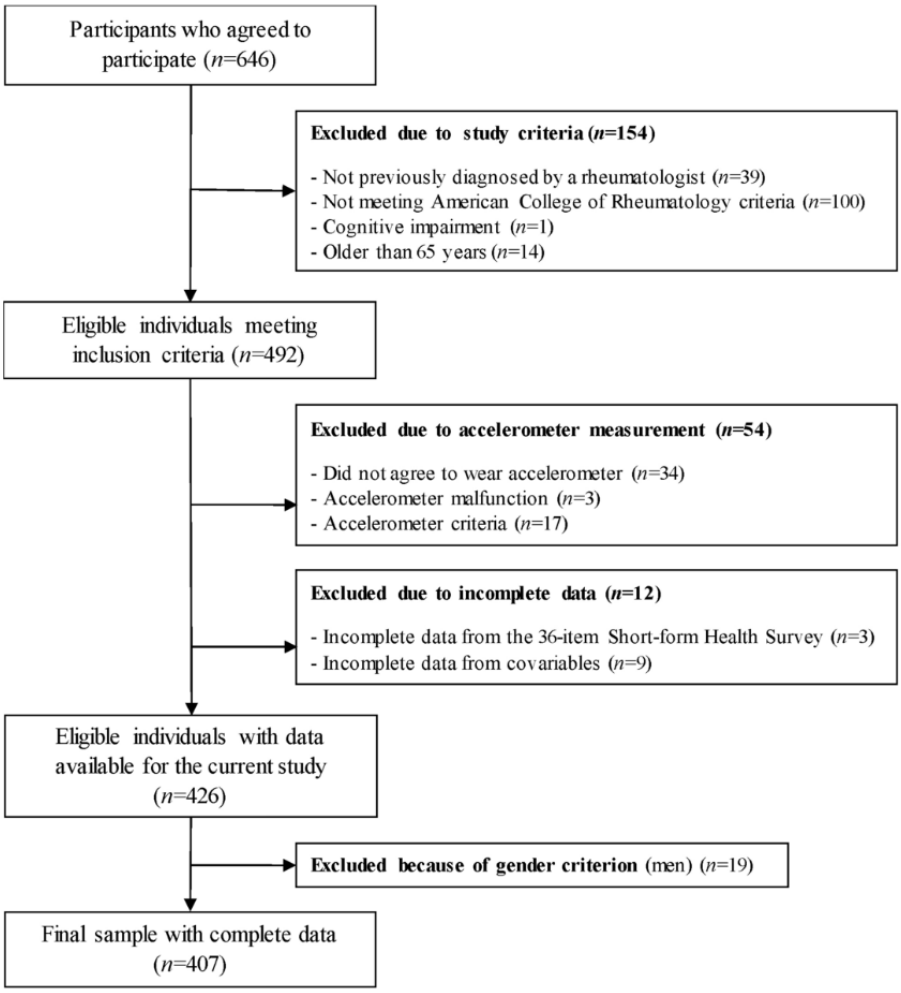
Therefore, we aimed to examine (1) the association of the patterns of ST (ie, ST accumulated in bouts and frequency of sedentary bouts) in different bout lengths (≥10 min, ≥20 min, ≥30 min, and ≥60 min) with the HRQoL in women with fibromyalgia, (2) the combined association of total ST and sedentary bout duration with HRQoL, and (3) whether these associations are independent of MVPA.

Methods

Recruitment

A representative sample of fibromyalgia patients from the south of Spain—Andalusia—was recruited for the al-Ándalus project via fibromyalgia associations, internet advertisement, flyers, and email. Written informed consent from all participants (N=646) was obtained. Inclusion criteria for this study require that participants (1) be previously diagnosed by a rheumatologist and meet the 1990 American College of Rheumatology fibromyalgia criteria [1], (2) do not have either acute or terminal illness or severe cognitive impairment, and (3) are 65 years of age or younger. The flowchart of participants included in this study is shown in Figure 1. The Ethics Committee of the Hospital Virgen de las Nieves, Granada, Spain, reviewed and approved the study.

Figure 1. Flow diagram of inclusion of women with fibromyalgia from the al-Ándalus project included in this study (N=407).



Measurements

Sedentary Time and Physical Activity

Patients wore the GT3X+ triaxial accelerometer (ActiGraph) on the hip for 9 days, 24 hours per day, except during water-based activities. Activity counts were measured at a rate of 30 Hz and stored at an epoch length of 1 minute [22]. Accelerometer-wearing time was obtained by subtracting the sleeping time and nonwear periods from each day. Sleeping time was obtained from a sleep diary, in which patients reported the time they went to bed and the time they woke up. Nonwear periods were obtained by applying Choi's algorithm [23]. Bouts of 90 continuous minutes of 0 counts were considered nonwear periods. To eliminate reactivity from the awareness of being monitored, we excluded PA data from the first day. The last day, when the device was returned, was excluded from the analysis as well. A total of 7 continuous days of recording, with a minimum of 10 valid hours per day, was the minimum criteria for being included in the study analysis.

ST and MVPA were calculated based upon recommended PA vector magnitude cut points [22,24]: 0-199 and ≥2690 counts per minute (cpm), respectively. A sedentary bout was defined as the number of consecutive minutes during which the accelerometer registered less than 200 cpm. Four sedentary bout-length categories were used in this study: ≥10 min, ≥20 min, ≥30 min, and ≥60 min. For each sedentary bout length, we obtained the following variables related to patterns of ST: (1) percentage of total ST accumulated in bouts (total time accumulated in bouts/total ST × 100) and (2) frequency of bouts (number of bouts/sedentary hours).

Data download, reduction, cleaning, and analyses were performed using ActiGraph's desktop software: ActiLife, version 6.11.7.

Health-Related Quality of Life

The HRQoL was assessed using the 36-item Short-Form Health Survey (SF-36) [25]. The SF-36 is composed of 36 items that assess eight dimensions of health (ie, physical functioning, physical role, bodily pain, general health, social functioning, emotional role, mental health, and vitality) and two component summary scores (ie, physical component summary [PCS] and mental component summary [MCS]). The score in each of the eight dimensions is standardized and ranges from 0 (*worst health status*) to 100 (*best health status*).

Sociodemographic and Clinical Data

We collected sociodemographic and clinical data using a self-reported questionnaire that included age, marital status (married/not married), education level (university/nonuniversity), and occupational status (working/not working). Patients also reported the consumption of

antidepressants and analgesics (yes/no) during the previous 2 weeks.

Anthropometry and Body Composition

Weight (kg) and total body fat (%) were measured using bioelectrical impedance with the InBody R20 (Biospace) body composition analyzer. Patients were asked neither to have a shower, practice intense PA, nor ingest large amounts of fluid and/or food in the 2 hours before the measurement. Patients removed their clothing and any metal objects from their bodies during the assessment.

Impact of the Disease

The Revised Fibromyalgia Impact Questionnaire (FIQR) [26] assesses overall fibromyalgia severity through a wide range of symptoms, comorbidities, and complaints related to this chronic condition. It is a self-administered questionnaire with 21 individual questions, with a rating scale of 0-10. The adjusted FIQR total score ranges from 0 to 100, with a higher score indicating greater impact of the syndrome on a person's life.

Statistical Analysis

Descriptive continuous data are shown as mean (SD), whereas categorical data are presented as n (%). To test the association between patterns of ST and HRQoL, we used linear regression analysis. The eight dimensions and the two summary components of the SF-36 were introduced as dependent variables in the models in separate regression models. Patterns of ST (ie, percentage of ST accumulated in bouts and frequency of bouts in all bout lengths) were introduced individually as predictor variables. Two types of models were built: (1) model 1 was controlled for age, total body fat percentage, occupational status, medication for pain and depression, and accelerometer wear time and (2) model 2 included model 1 plus MVPA.

The combined association of total ST and prolonged sedentary bout duration with HRQoL was studied through analyses of covariance. The subject pool was divided into four groups according to the median value of total ST (3216 min/week) and the median value of sedentary bout durations of 30 continuous minutes or more (47.7 min). A minimum duration of 30 continuous minutes was used to define prolonged ST following the criteria of previous studies [27]. The four groups created were (1) low total ST (≤ the median value) + low sedentary bout duration (≤ the median value), (2) low total ST + high sedentary bout duration (> the median value), (3) high total ST (> the median value) + low sedentary bout duration, and (4) high total ST + high sedentary bout duration. The analyses were adjusted for age, total body fat percentage, occupational status, medication for pain and depression, and accelerometer-wear time. Additional analyses including MVPA as covariate were performed.

For analyses, we used IBM SPSS Statistics for Windows, version 20.0 (IBM Corp). The statistical significance was set at $P<.05$.

Data Exclusion

The final sample size included in the analyses comprised 407 women with fibromyalgia. The flow diagram of women with fibromyalgia included in this study is shown in Figure 1.

Results

Table 1 provides an overview of the patients' sociodemographic and clinical characteristics. Table 2 includes the information related to PA and ST pattern characteristics (ie, percentage of total ST and frequency of bouts) in different bout lengths.

The association of the percentage of ST accumulated in bouts of different lengths with the SF-36 domains are shown in Table 3. Greater percentages of ST spent in all bout lengths were associated with worse physical function, bodily pain, vitality, and social function domains and the PCS (beta from -.20 to -.10, all $P<.05$). In addition, a higher percentage of ST spent in bouts of 60 minutes or more was related to a worsened physical role (beta=-.10, $P=.04$). Overall, these associations were independent of MVPA (all $P<.05$), except for the bodily pain (for bouts ≥10, ≥20, or ≥30 min) and physical role domains.

Table 1. Sociodemographic and clinical characteristics of the study participants (N=407).

Variables	Mean (SD) or n (%)
Age (years), mean (SD)	51.4 (7.6)
Algometer score (18-144), mean (SD)	43.2 (13.4)
Body mass index (kg/m^2), mean (SD)	28.4 (5.4)
Total body fat (%), mean (SD)	40.1 (7.6)
FIQR[a] score (0-100), mean (SD)	64.4 (16.7)
Health-related quality of life, SF-36 [b] score (0-100), mean (SD)	
Physical function	39.2 (18.9)
Physical role	33.2 (21.2)
Bodily pain	21.2 (14.7)
General health	28.5 (15.3)
Vitality	22.3 (17.7)
Social functioning	43.7 (24.7)
Emotional role	56.9 (27.9)
Mental health	46.2 (19.7)
Physical component	29.5 (6.9)
Mental component	36.0 (11.6)
Marital status, n (%)	
Married	311 (76.4)
Not married	96 (23.6)
Education level, n (%)	
Nonuniversity	349 (85.7)
University	58 (14.3)
Current occupational status, n (%)	
Working	107 (26.3)
Not working	300 (73.7)
Drug consumption, n (%)	
Analgesics	367 (90.2)
Antidepressants	232 (57.0)

[a]FIQR: Revised Fibromyalgia Impact Questionnaire.

[b]SF-36: 36-item Short-Form Health Survey.

Table 2. Sedentary patterns and physical activity (PA) variables of the study participants (N=407).

Sedentary behavior and PA	Mean (SD)
Accelerometer-wear time	923.0 (78.9)
Sedentary time (ST)	
Minutes per day	460.1 (104.1)
Percentage of wear time	49.9 (10.6)
Light PA	
Minutes per day	418.6 (91.8)
Percentage of wear time	45.3 (9.1)
Moderate PA	
Minutes per day	43.9 (29.5)
Percentage of wear time	4.8 (3.2)
Vigorous PA	
Minutes per day	0.4 (2.0)
Percentage of wear time	0.1 (0.2)
Moderate-to-vigorous PA	
Minutes per day	44.3 (30.1)
Percentage of wear time	4.8 (3.2)
Patterns of ST of different bout lengths	
≥10-minute bout	
Percentage of total ST accumulated (%)	59.2 (11.2)
Frequency of bouts (number of bouts/week)	83.7 (25.6)
≥20-minute bout	
Percentage of total ST accumulated (%)	38.5 (12.8)
Frequency of bouts (number of bouts/week)	34.3 (14.6)
≥30-minute bout	
Percentage of total ST accumulated (%)	26.7 (12.3)
Frequency of bouts (number of bouts/week)	17.9 (9.6)
≥60-minute bout	
Percentage of total ST accumulated (%)	10.3 (8.9)
Frequency of bouts (number of bouts/week)	4.3 (3.7)

Table 3. Association of the percentage of sedentary time (ST) accumulated in bouts of different lengths with 36-item Short-Form Health Survey (SF-36) dimensions (N=407).

Dimensions and models	Percentage of ST accumulated in bouts of different lengths															
	≥10-minute bout				≥20-minute bout				≥30-minute bout				≥60-minute bout			
	B[a]	SE	Beta[b]	*P*	B	SE	Beta	*P*	B	SE	Beta	*P*	B	SE	Beta	*P*
Physical function																
Model 1[c]	-0.267	0.086	-.159	.002	-0.253	0.074	-.171	.001	-0.271	0.077	-.176	<.001	-0.428	0.104	-.201	<.001
Model 2[d]	-0.226	0.089	-.134	.01	-0.221	0.077	-.149	.004	-0.239	0.079	-.156	.002	-0.396	0.105	-.186	<.001
Physical role																
Model 1	-0.126	0.094	-.067	.18	-0.137	0.081	-.082	.09	-0.159	0.084	-.092	.06	-0.239	0.114	-.100	.04
Model 2	-0.063	0.097	-.033	.52	-0.089	0.083	-.054	.29	-0.115	0.086	-.067	.18	-0.195	0.115	-.082	.09
Bodily pain																
Model 1	-0.130	0.065	-.099	.045	-0.108	0.056	-.094	.05	-0.114	0.058	-.096	.048	-0.190	0.078	-.115	.02
Model 2	-0.106	0.067	-.081	.12	-0.089	0.058	-.077	.13	-0.096	0.059	-.080	.11	-0.171	0.079	-.104	.03
General health																
Model 1	-0.034	0.069	-.025	.62	-0.048	0.060	-.040	.42	-0.058	0.062	-.047	.35	-0.076	0.084	-.044	.37
Model 2	-0.029	0.072	-.021	.69	-0.046	0.062	-.038	.46	-0.056	0.064	-.045	.38	-0.073	0.085	-.042	.39
Vitality																
Model 1	-0.252	0.080	-.160	.002	-0.204	0.069	-.148	.003	-0.204	0.072	-.142	.004	-0.278	0.097	-.140	.004
Model 2	-0.209	0.083	-.133	.01	-0.168	0.071	-.122	.02	-0.169	0.073	-.117	.02	-0.242	0.098	-.122	.01
Social functioning																
Model 1	-0.399	0.106	-.181	<.001	-0.351	0.091	-.182	<.001	-0.361	0.095	-.180	<.001	-0.500	0.129	-.181	<.001
Model 2	-0.285	0.108	-.130	.01	-0.261	0.093	-.135	.01	-0.275	0.095	-.137	.004	-0.412	0.128	-.149	.001
Emotional role																
Model 1	0.040	0.121	.016	.74	0.012	0.105	.005	.91	-0.016	0.109	-.007	.88	-0.034	0.148	-.011	.82
Model 2	0.077	0.126	.031	.54	0.039	0.109	.018	.72	0.008	0.112	.004	.94	-0.010	0.150	-.003	.95
Mental health																
Model 1	0.028	0.086	.016	.74	-0.015	0.075	-.010	.84	-0.029	0.077	-.018	.71	-0.104	0.105	-.047	.32
Model 2	0.059	0.090	.034	.51	0.006	0.077	.004	.94	-0.010	0.079	-.006	.90	-0.087	0.106	-.039	.41
Physical component																
Model 1	-0.096	0.032	-.156	.003	-0.085	0.027	-.158	.002	-0.089	0.028	-.159	.002	-0.130	0.038	-.168	.001

Dimensions and models	Percentage of ST accumulated in bouts of different lengths															
	≥10-minute bout				≥20-minute bout				≥30-minute bout				≥60-minute bout			
	B[a]	SE	Beta[b]	*P*	B	SE	Beta	*P*	B	SE	Beta	*P*	B	SE	Beta	*P*
Model 2	-0.083	0.033	-.135	.01	-0.075	0.028	-.139	.01	-0.079	0.029	-.141	.01	-0.119	0.039	-.154	.002
Mental component																
Model 1	-0.023	0.050	-.022	.64	-0.031	0.043	-.034	.47	-0.039	0.045	-.041	.39	-0.065	0.061	-.050	.28
Model 2	0.003	0.052	.003	.95	-0.011	0.044	-.012	.80	-0.020	0.046	-.021	.66	-0.047	0.061	-.036	.44

[a]B: nonstandardized regression coefficient. Linear regression models were built using *Enter* method, with SF-36 domains as dependent variables and percentage of ST in different bout lengths as independent variables.

[b]Beta: standardized regression coefficient.

[c]Model 1: adjusted for age, fat percentage, occupational status, medication for pain, medication for depression, and accelerometer wear time.

[d]Model 2: analysis controlled for model 1 + moderate-to-vigorous PA.

Table 4 shows the association of the frequency of bouts of ST of different lengths with the SF-36 domains. A higher frequency of sedentary bouts 20 minutes or longer was associated with worsened vitality and social function (beta=-.12 and -.13, respectively, all *P*<.05). A higher frequency of sedentary bouts 30 minutes or longer was associated with worsened physical function, vitality, social function, and PCS scores (beta from -.15 to -.12, all *P*<.05). A higher frequency of sedentary bouts 60 minutes or longer was associated with worsened physical function, physical role, bodily pain, vitality, social function, and PCS scores (beta from -.19 to -.10, all *P*<.05). These associations were independent of MVPA, except for the association with physical role, vitality, and social function in bouts 20 minutes or longer.

Figure 2 shows the combined association of total ST and sedentary bout duration with the SF-36 domains, the PCS, and the MCS. Participants with low total ST and low sedentary bout duration presented better physical function (mean difference 8.73 units, 95% CI 2.31-15.15), social function (mean difference 10.51 units, 95% CI 2.59-18.44), and PCS (mean difference 2.71 units, 95% CI 0.36-5.06) compared to participants with high total ST and high sedentary bout duration (all *P*<.02). Additional analyses showed that the differences in the physical function (*P*=.045) were independent of MVPA.

Table 4. Association of the frequency of bouts of sedentary time (ST) of different lengths with 36-item Short-Form Health Survey (SF-36) dimensions (N=407).

Dimensions and models	Frequency of bouts (number of bouts/sedentary hours) of different bout lengths															
	≥10-minute bout				≥20-minute bout				≥30-minute bout				≥60-minute bout			
	B[a]	SE	Beta[b]	*P*	B	SE	Beta	*P*	B	SE	Beta	*P*	B	SE	Beta	*P*
Physical function																
Model 1[c]	2.386	4.237	.028	.57	-10.754	5.882	-.092	.07	-18.793	7.663	-.123	.02	-61.611	16.061	-.188	<.001
Model 2[d]	3.728	4.240	.044	.38	-7.970	6.002	-.068	.19	-15.416	7.820	-.101	.049	-56.713	16.208	-.173	.001
Physical role																
Model 1	3.602	4.575	.038	.43	-4.921	6.375	-.038	.44	-12.379	8.315	-.072	.14	-35.905	17.572	-.098	.04
Model 2	5.161	4.572	.055	.26	-1.341	6.490	-.010	.84	-8.033	8.469	-.047	.34	-29.379	17.693	-.080	.10
Bodily pain																
Model 1	-1.811	3.163	-.028	.57	-6.140	4.399	-.068	.16	-9.957	5.741	-.084	.08	-31.088	12.109	-.122	.01
Model 2	-1.128	3.179	-.017	.72	-4.636	4.500	-.051	.30	-8.101	5.873	-.068	.17	-28.355	12.245	-.111	.02
General health																
Model 1	1.978	3.370	.029	.56	-2.700	4.696	-.029	.57	-6.395	6.132	-.052	.30	-14.374	12.987	-.054	.27
Model 2	2.167	3.399	.032	.52	-2.433	4.818	-.026	.61	-6.189	6.289	-.050	.33	-13.956	13.166	-.053	.29
Vitality																
Model 1	-4.226	3.928	-.053	.28	-13.300	5.441	-.122	.02	-19.025	7.101	-.133	.01	-44.261	15.015	-.145	.003
Model 2	-2.976	3.931	-.038	.45	-10.617	5.549	-.097	.06	-15.624	7.240	-.109	.03	-38.871	15.125	-.127	.01
Social functioning																
Model 1	-2.477	5.242	-.022	.64	-19.997	7.237	-.132	.01	-30.153	9.429	-.151	.001	-73.977	19.887	-.173	<.001
Model 2	0.493	5.150	.004	.92	-13.339	7.267	-.088	.07	-21.717	9.472	-.109	.02	-60.962	19.730	-.143	.002
Emotional role																
Model 1	4.353	5.913	.035	.46	1.973	8.246	.011	.81	-2.809	10.777	-.012	.79	-4.716	22.830	-.010	.84
Model 2	5.130	5.958	.041	.39	3.877	8.449	.023	.65	-0.610	11.041	-.003	.96	-1.328	23.120	-.003	.95
Mental health																
Model 1	5.702	4.196	.065	.18	0.485	5.861	.004	.93	-2.013	7.659	-.013	.79	-19.328	16.197	-.057	.23
Model 2	6.372	4.225	.072	.13	1.989	6.003	.016	.74	-0.195	7.844	-.001	.98	-16.941	16.404	-.050	.30
Physical component																
Model 1	-0.238	1.557	-.008	.88	-4.104	2.160	-.097	.06	-6.827	2.816	-.122	.02	-19.387	5.930	-.162	.001

Dimensions and models	Frequency of bouts (number of bouts/sedentary hours) of different bout lengths															
	≥10-minute bout				≥20-minute bout				≥30-minute bout				≥60-minute bout			
	B[a]	SE	Beta[b]	*P*	B	SE	Beta	*P*	B	SE	Beta	*P*	B	SE	Beta	*P*
Model 2	0.186	1.562	.006	.91	-3.222	2.208	-.076	.15	-5.754	2.877	-.103	.046	-17.763	5.989	-.149	.003
Mental component																
Model 1	1.309	2.431	.025	.59	-1.825	3.387	-.025	.59	-4.017	4.424	-.043	.36	-11.036	9.365	-.055	.24
Model 2	1.910	2.440	.037	.43	-0.438	3.461	-.006	.90	-2.300	4.520	-.024	.61	-8.442	9.459	-.042	.37

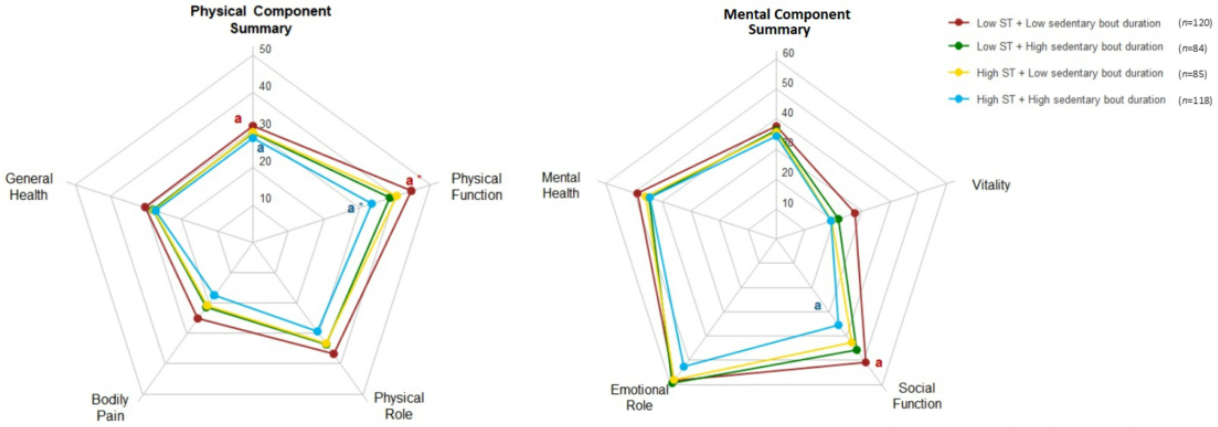
[a]B: nonstandardized regression coefficient. Linear regression models built using *Enter* method, with SF-36 domains as dependent variables and percentage of ST in different bout lengths as independent variables.

[b]Beta: standardized regression coefficient.

[c]Model 1: adjusted for age, fat percentage, occupational status, medication for pain, medication for depression, accelerometer wear time, and ST.

[d]Model 2: analysis controlled for model 1 + moderate-to-vigorous PA.

Figure 2. Combined association of total sedentary time (ST) and prolonged sedentary bouts of at least 30 minutes with health-related quality of life. Estimated means represent values after adjustment for age, total body fat percentage, occupational status, medication for pain, medication for depression, and accelerometer-wear time. Common superscripts indicate significant ($P \leq .05$) differences between groups with the same letter when adjusting for age, body fat percentage, occupational status, medication for pain, medication for depression, and accelerometer-wear time. Asterisks represent significant differences for additional adjustment for moderate-to-vigorous physical activity ($P = .045$).



Discussion

Principal Findings

The main findings of this study suggest that higher percentages of ST spent in different bout lengths were associated with worsened HRQoL, including physical function, bodily pain, vitality, and social function domains, as well as the PCS. Also, higher frequencies of sedentary bouts were associated with worsened HRQoL, including physical function, bodily pain, vitality, and social function domains, as well as the PCS, especially in longer bout durations. Patients characterized by high total ST and high sedentary bout duration presented worsened physical function, social function, and PCS scores. These associations were generally independent of the MVPA performed for long bout lengths. These findings entail a first step toward the understanding of free-living sedentary behavior and its association with HRQoL in fibromyalgia. This supports the implementation of mHealth devices, which allow self-monitoring and immediate feedback of daily living behaviors to patients. Future studies might determine whether this approach is successful by reducing prolonged ST in this population.

Limitations

This study has several limitations that should be underlined. Because our results were derived from a cross-sectional design, the associations cannot be explained via a causal pathway. In addition, due to the large quantity of factors related to HRQoL, it is difficult to ascertain the true nature of the association found between variables. Because only women took part in this study, future studies should investigate whether these associations might extend to men as well. Among its strengths, this study includes a relatively large sample size of women with

fibromyalgia representative from the south of Spain (ie, Andalusia). According to a recent study, measurement of the actual dose of exercise and daily mobility are essential to establish relationships of these behaviors with health [21]. In this sense, ST and PA were objectively assessed in this study through a wearable tool that enabled researchers to monitor the type, quantity, and quality of everyday activities of patients via accelerometry, which is considered a more reliable technique than questionnaires in the study of fibromyalgia [28]. Future intervention studies with mHealth devices that incorporate in situ information are warranted in this population to ascertain whether fibromyalgia patients change their sedentary behaviors.

Comparison With Prior Work

To date, most of the previous research on ST and health in fibromyalgia has been limited to the study of total ST [7,29]. In addition, few studies have objectively characterized ST through accelerometry in these patients [18,30] and only one of them [18] reported the values of sustained ST (>1 hour). Ellingson et al demonstrated that sustained ST (>60 min) was associated with worse pain modulation in fibromyalgia—assessed through magnetic resonance imaging—to a greater extent compared to total ST [18]. Congruently, this study showed negative associations between time spent in sedentary bouts (≥10 min, ≥20 min, ≥30 min, and ≥60 min) and the SF-36 body pain dimension. Therefore, we extend the connection between sustained ST and pain to patient-reported instruments. Also, the interruptions of these sedentary bouts might be relevant for pain in this population, given that frequency of sedentary bouts (≥60 min) was negatively associated with bodily pain scores. Following the findings by Ellingson et al, increased pain in sustained ST could be due to the impaired activity in the prefrontal cortices and sensory regions (ie, pre- and postcentral gyri) of these patients [18]. Because the bodily pain domain of the SF-36 not only encompasses objective levels of pain but also the perceived limitations due to it, the contribution of other factors influencing patients' perceptions, such as self-efficacy or pain coping strategies [31], could also take part in this relationship.

Although the influence of patterns of ST has not been explored in fibromyalgia, a direct relationship between increased total ST and fatigue has been described [29]. In agreement with our findings, one previous study in healthy women showed that prolonged ST accumulated in bouts of at least 1 hour were negatively associated with vitality scores of the SF-36 and other fatigue-related variables [32]. Despite the cross-sectional design of these findings that precludes the causal explanation, other experimental studies observed increases in fatigue levels during uninterrupted sitting in adults with overweight and obese status [33] and type 2 diabetes [34], or decreases in fatigue as a result of reducing prolonged sitting [33]. The relationship between ST and fatigue might be explained through physiological, psychological, and social factors that contribute to this multifaceted phenomenon. For instance, prolonged ST could alter the sympathetic nervous system (ie, through a lower heart rate, decreased plasma level of dihydroxyphenylalanine, and increased plasma level of dihydroxyphenylglycol) [33], promote muscle fatigue through sustained activation of low-threshold motor units in sedentary positions [35], or negatively influence sleep quality [7].

In fibromyalgia, there is also a gap in the literature regarding the influence of ST and its patterns on social limitations due to health. For other social-related aspects, Soursa et al stated that patients with fibromyalgia with the lowest PA levels and, presumably, higher levels of ST, had fewer social interactions compared to those doing more PA [36]. No evidence is available on how patterns of ST could influence social function in other populations either, yet interpersonal factors (eg, family, friends, and social networks) are well-known determinants of sedentary behaviors [37]. The passive nature of different sedentary activities (eg, watching television or sitting at the computer) that are accompanied with decreased communication [38] could also lead to poor social networking and participation [39]. Therefore, future research might ascertain whether breaking prolonged ST could positively influence this construct of health (eg, through an increased opportunity to interact with others) or whether strategies aimed at increasing social support may lead to more favorable patterns of accumulation of ST.

To our knowledge, no previous studies have linked patterns of ST to physical function in fibromyalgia. The physical function domain assesses activities of daily living (eg, bathing, dressing, walking several blocks, and lifting or carrying groceries) that typically require a combination of flexibility, strength, and cardiorespiratory fitness, which are related to HRQoL [40]. Previous evidence in adults or older adults showed a decreased physical function, assessed through physical fitness tests, in relation to more deleterious patterns of device-measured ST, such as reduced breaks in ST [41,42], increased sedentary bout duration [41], or increased total prolonged ST [42]. Sedentary periods are linked to skeletal muscle inactivity [43] and are thought to accelerate sarcopenia and loss of aerobic capacity [44], which could negatively affect physical function. Therefore, increases in physical function could be optimized by avoiding the accumulation of ST in prolonged periods and reducing the duration of these ST periods, which needs to be confirmed in future intervention studies.

We observed that, overall, the strength of the associations between the patterns of ST and HRQoL was reduced but still significant when considering MVPA. This finding is congruent with a recent meta-analysis concluding that the deleterious health effects associated with ST generally decrease in magnitude among people with higher levels of PA [13]. Our results also showed that, for certain patterns of ST in shorter bout lengths (<60 min), the associations with HRQoL were not significant anymore when considering MVPA. Therefore, performing MVPA could have a protective effect only when ST is accrued in low-duration bouts and could be especially relevant for certain domains of HRQoL, such as bodily pain, physical role, vitality, or social function. Interestingly, meeting the current guidelines of MVPA in bouts of at least 10 minutes was found to neutralize the negative association of prolonged ST with fatigue in healthy women [32]. Hence, it is possible that the patterns of accumulation of MVPA could also influence the capacity of this behavior of counteracting the negative effects of prolonged ST.

Conclusions

In conclusion, our findings indicate that higher ST spent in diverse bout lengths and a higher frequency of sedentary bouts, especially in longer bout durations, is associated with worsened HRQoL, more specifically with physical function, bodily pain, vitality, and social function domains, as well as the PCS. Patients that are highly sedentary and present longer sedentary bout durations have worsened physical function, social function, and PCS scores. Although these associations were generally independent of MVPA in long sedentary bout lengths, this intensity of PA could play a positive role when ST is accumulated in shorter bouts. Future intervention studies using mHealth devices that incorporate immediate feedback for users are warranted in this population to ascertain whether fibromyalgia patients change their sedentary behaviors.

Acknowledgments

We thank the collaborators of the al-Ándalus project and all the members of the Physical Activity for Health Promotion (PAHELP; CTS-1018) research group. We also gratefully acknowledge all the study participants for their collaboration. This study was supported by the Spanish Ministries of Economy and Competitiveness (I+D+i DEP2010-15639; I+D+i DEP2013-40908-R, BES-2014-067612) and the Spanish Ministry of Education (FPU 15/00002). This study was also partially funded by the University of Granada, Plan Propio de Investigación 2016, Excellence actions: Units of Excellence; Unit of Excellence on Exercise and Health (UCEES), by the Junta de Andalucía, Consejería de Conocimiento, Investigación y Universidades and European Regional Development Fund (ERDF), ref SOMM17/6107/UGR, and by the University of Cádiz (Research and Transfer Activity Promotion Program). This study is part of a PhD thesis of the Biomedicine Program of the University of Granada.

Conflicts of Interest

None declared.

References

1. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia: Report of the Multicenter Criteria Committee. Arthritis Rheum 1990 Feb;33(2):160-172. [doi: 10.1002/art.1780330203] [Medline: 2306288]
2. McLoughlin MJ, Colbert LH, Stegner AJ, Cook DB. Are women with fibromyalgia less physically active than healthy women? Med Sci Sports Exerc 2011 May;43(5):905-912 [FREE Full text] [doi: 10.1249/MSS.0b013e3181fca1ea] [Medline: 20881881]
3. Björnsdóttir SV, Jónsson SH, Valdimarsdóttir UA. Functional limitations and physical symptoms of individuals with chronic pain. Scand J Rheumatol 2013;42(1):59-70. [doi: 10.3109/03009742.2012.697916] [Medline: 23126682]
4. Busch AJ, Webber SC, Brachaniec M, Bidonde J, Bello-Haas VD, Danyliw AD, et al. Exercise therapy for fibromyalgia. Curr Pain Headache Rep 2011 Oct;15(5):358-367 [FREE Full text] [doi: 10.1007/s11916-011-0214-2] [Medline: 21725900]
5. Ellingson LD, Shields MR, Stegner AJ, Cook DB. Physical activity, sustained sedentary behavior, and pain modulation in women with fibromyalgia. J Pain 2012 Feb;13(2):195-206 [FREE Full text] [doi: 10.1016/j.jpain.2011.11.001] [Medline: 22245361]
6. Soriano-Maldonado A, Ruiz JR, Aparicio VA, Estévez-López F, Segura-Jiménez V, Álvarez-Gallardo IC, et al. Association of physical fitness with pain in women with fibromyalgia: The al-Ándalus project. Arthritis Care Res (Hoboken) 2015 Nov;67(11):1561-1570 [FREE Full text] [doi: 10.1002/acr.22610] [Medline: 25939406]
7. Borges-Cosic M, Aparicio V, Estévez-López F, Soriano-Maldonado A, Acosta-Manzano P, Gavilán-Carrera B, et al. Sedentary time, physical activity, and sleep quality in fibromyalgia: The al-Ándalus project. Scand J Med Sci Sports 2019 Feb;29(2):266-274. [doi: 10.1111/sms.13318] [Medline: 30306645]
8. Gavilán-Carrera B, Segura-Jiménez V, Estévez-López F, Álvarez-Gallardo IC, Soriano-Maldonado A, Borges-Cosic M, al-Ándalus Project. Association of objectively measured physical activity and sedentary time with health-related quality of life in women with fibromyalgia: The al-Ándalus project. J Sport Health Sci 2019 May;8(3):258-266 [FREE Full text] [doi: 10.1016/j.jshs.2018.07.001] [Medline: 31193271]
9. Patterson R, McNamara E, Tainio M, de Sá TH, Smith A, Sharp S, et al. Sedentary behaviour and risk of all-cause, cardiovascular and cancer mortality, and incident type 2 diabetes: A systematic review and dose response meta-analysis. Eur J Epidemiol 2018 Sep;33(9):811-829 [FREE Full text] [doi: 10.1007/s10654-018-0380-1] [Medline: 29589226]
10. Owen N, Sparling PB, Healy GN, Dunstan DW, Matthews CE. Sedentary behavior: Emerging evidence for a new health risk. Mayo Clin Proc 2010 Dec;85(12):1138-1141 [FREE Full text] [doi: 10.4065/mcp.2010.0444] [Medline: 21123641]
11. Segura-Jiménez V, Álvarez-Gallardo IC, Estévez-López F, Soriano-Maldonado A, Delgado-Fernández M, Ortega FB, et al. Differences in sedentary time and physical activity between female patients with fibromyalgia and healthy controls: The al-Ándalus project. Arthritis Rheumatol 2015 Nov;67(11):3047-3057 [FREE Full text] [doi: 10.1002/art.39252] [Medline: 26108350]

12. Ekelund U, Steene-Johannessen J, Brown WJ, Fagerland MW, Owen N, Powell KE, et al. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. Lancet 2016 Sep;388(10051):1302-1310. [doi: 10.1016/s0140-6736(16)30370-1]
13. Biswas A, Oh P, Faulkner G, Bajaj R, Silver M, Mitchell M, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: A systematic review and meta-analysis. Ann Intern Med 2015 Jan 20;162(2):123-132. [doi: 10.7326/M14-1651] [Medline: 25599350]
14. Macfarlane GJ, Kronisch C, Dean LE, Atzeni F, Häuser W, Fluß E, et al. EULAR revised recommendations for the management of fibromyalgia. Ann Rheum Dis 2017 Feb;76(2):318-328. [doi: 10.1136/annrheumdis-2016-209724] [Medline: 27377815]
15. Peddie M, Bone J, Rehrer N, Skeaff C, Gray A, Perry T. Breaking prolonged sitting reduces postprandial glycemia in healthy, normal-weight adults: A randomized crossover trial. Am J Clin Nutr 2013 Aug;98(2):358-366. [doi: 10.3945/ajcn.112.051763] [Medline: 23803893]
16. Howard BJ, Fraser SF, Sethi P, Cerin E, Hamilton MT, Owen N, et al. Impact on hemostatic parameters of interrupting sitting with intermittent activity. Med Sci Sports Exerc 2013 Jul;45(7):1285-1291. [doi: 10.1249/MSS.0b013e318285f57e] [Medline: 23439415]
17. Dunstan DW, Kingwell BA, Larsen R, Healy GN, Cerin E, Hamilton MT, et al. Breaking up prolonged sitting reduces postprandial glucose and insulin responses. Diabetes Care 2012 May;35(5):976-983 [FREE Full text] [doi: 10.2337/dc11-1931] [Medline: 22374636]
18. Ellingson LD, Shields MR, Stegner AJ, Cook DB. Physical activity, sustained sedentary behavior, and pain modulation in women with fibromyalgia. J Pain 2012 Feb;13(2):195-206 [FREE Full text] [doi: 10.1016/j.jpain.2011.11.001] [Medline: 22245361]
19. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Breaks in sedentary time: Beneficial associations with metabolic risk. Diabetes Care 2008 Apr;31(4):661-666. [doi: 10.2337/dc07-2046] [Medline: 18252901]
20. 2018 Physical Activity Guidelines Advisory Committee. 2018 Physical Activity Guidelines Advisory Committee Scientific Report. Washington, DC: US Department of Health and Human Services; 2018. URL: https://health.gov/paguidelines/second-edition/report/pdf/PAG_Advisory_Committee_Report.pdf [accessed 2020-02-01]
21. Dobkin B, 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]
22. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. J Sci Med Sport 2011 Sep;14(5):411-416. [doi: 10.1016/j.jsams.2011.04.003] [Medline: 21616714]
23. Choi L, Ward SC, Schnelle JF, Buchowski MS. Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. Med Sci Sports Exerc 2012 Oct;44(10):2009-2016 [FREE Full text] [doi: 10.1249/MSS.0b013e318258cb36] [Medline: 22525772]
24. Aguilar-Farías N, Brown WJ, Peeters GM. ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. J Sci Med Sport 2014 May;17(3):293-299. [doi: 10.1016/j.jsams.2013.07.002] [Medline: 23932934]
25. Alonso J, Prieto L, Antó JM. The Spanish version of the SF-36 Health Survey (the SF-36 health questionnaire): an instrument for measuring clinical results [Article in Spanish]. Med Clin (Barc) 1995 May 27;104(20):771-776. [Medline: 7783470]
26. Rivera J, Gonzalez T. The Fibromyalgia Impact Questionnaire: A validated Spanish version to assess the health status in women with fibromyalgia. Clin Exp Rheumatol 2004;22(5):554-560. [doi: 10.1007/springerreference_184575]
27. Diaz KM, Goldsmith J, Greenlee H, Strizich G, Qi Q, Mossavar-Rahmani Y, et al. Prolonged, uninterrupted sedentary behavior and glycemic biomarkers among us Hispanic/Latino adults. Circulation 2017 Oct 10;136(15):1362-1373. [doi: 10.1161/circulationaha.116.026858]
28. Segura-Jiménez V, Alvarez-Gallardo IC, Romero-Zurita A, Camiletti-Moirón D, Munguía-Izquierdo D, Carbonell-Baeza A, et al. Comparison of physical activity using questionnaires (leisure time physical activity instrument and physical activity at home and work instrument) and accelerometry in fibromyalgia patients: The Al-Ándalus project. Arch Phys Med Rehabil 2014 Oct;95(10):1903-1911.e2. [doi: 10.1016/j.apmr.2014.05.015] [Medline: 24909586]
29. Segura-Jiménez V, Borges-Cosic M, Soriano-Maldonado A, Estévez-López F, Álvarez-Gallardo IC, Herrador-Colmenero M, et al. Association of sedentary time and physical activity with pain, fatigue, and impact of fibromyalgia: The al-Ándalus study. Scand J Med Sci Sports 2017 Jan;27(1):83-92. [doi: 10.1111/sms.12630] [Medline: 26644186]
30. Bernard P, Hains-Monfette G, Atoui S, Kingsbury C. Differences in daily objective physical activity and sedentary time between women with self-reported fibromyalgia and controls: Results from the Canadian health measures survey. Clin Rheumatol 2018 Aug;37(8):2285-2290. [doi: 10.1007/s10067-018-4139-6] [Medline: 29754181]
31. Lee J, Lee K, Park D, Kim S, Nah S, Lee JH, et al. Determinants of quality of life in patients with fibromyalgia: A structural equation modeling approach. PLoS One 2017;12(2):e0171186 [FREE Full text] [doi: 10.1371/journal.pone.0171186] [Medline: 28158289]
32. Ellingson LD, Kuffel AE, Vack NJ, Cook DB. Active and sedentary behaviors influence feelings of energy and fatigue in women. Med Sci Sports Exerc 2014 Jan;46(1):192-200. [doi: 10.1249/MSS.0b013e3182a036ab] [Medline: 23783259]

33. Wennberg P, Boraxbekk C, Wheeler M, Howard B, Dempsey PC, Lambert G, et al. Acute effects of breaking up prolonged sitting on fatigue and cognition: A pilot study. BMJ Open 2016 Feb 26;6(2):e009630 [FREE Full text] [doi: 10.1136/bmjopen-2015-009630] [Medline: 26920441]
34. Dempsey PC, Dunstan DW, Larsen RN, Lambert GW, Kingwell BA, Owen N. Prolonged uninterrupted sitting increases fatigue in type 2 diabetes. Diabetes Res Clin Pract 2018 Jan;135:128-133. [doi: 10.1016/j.diabres.2017.11.001] [Medline: 29129482]
35. Hagg G. Static work loads and occupational myalgia: A new explanation model. In: Proceedings of the 8th Congress of the International Society of Electrophysiological Kinesiology. 1991 Presented at: 8th Congress of the International Society of Electrophysiological Kinesiology; August 12-16, 1990; Baltimore, MD p. 141-144.
36. Suorsa K, Lynch-Jordan A, Tran S, Edwards N, Kashikar-Zuck S. Rates of physical activity and perceived social support among young adult women with juvenile-onset fibromyalgia. J Pain 2016 Apr;17(4):S100-S101. [doi: 10.1016/j.jpain.2016.01.311]
37. Chastin SF, De Craemer M, Lien N, Bernaards C, Buck C, Oppert J, DEDIPAC Consortium, Expert Working Group and Consensus Panel. The SOS-framework (Systems of Sedentary behaviours): An international transdisciplinary consensus framework for the study of determinants, research priorities and policy on sedentary behaviour across the life course: A DEDIPAC study. Int J Behav Nutr Phys Act 2016 Jul 15;13:83 [FREE Full text] [doi: 10.1186/s12966-016-0409-3] [Medline: 27421750]
38. Kraut R, Patterson M, Lundmark V, Kiesler S, Mukophadhyay T, Scherlis W. Internet paradox: A social technology that reduces social involvement and psychological well-being? Am Psychol 1998;53(9):1017-1031. [doi: 10.1037//0003-066x.53.9.1017]
39. Kikuchi H, Inoue S, Fukushima N, Takamiya T, Odagiri Y, Ohya Y, et al. Social participation among older adults not engaged in full- or part-time work is associated with more physical activity and less sedentary time. Geriatr Gerontol Int 2017 Nov;17(11):1921-1927. [doi: 10.1111/ggi.12995] [Medline: 28230301]
40. Álvarez-Gallardo IC, Soriano-Maldonado A, Segura-Jiménez V, Estévez-López F, Camiletti-Moirón D, Aparicio V, et al. High levels of physical fitness are associated with better health-related quality of life in women with fibromyalgia: The al-Ándalus project. Phys Ther 2019 Nov 25;99(11):1481-1494. [doi: 10.1093/ptj/pzz113] [Medline: 31392995]
41. van der Velde JH, Savelberg HH, van der Berg JD, Sep SJ, van der Kallen CJ, Dagnelie PC, et al. Sedentary behavior is only marginally associated with physical function in adults aged 40-75 years: The Maastricht study. Front Physiol 2017;8:242 [FREE Full text] [doi: 10.3389/fphys.2017.00242] [Medline: 28487660]
42. Dogra S, Clarke JM, Copeland JL. Prolonged sedentary time and physical fitness among Canadian men and women aged 60 to 69. Health Rep 2017 Feb 15;28(2):3-9 [FREE Full text] [Medline: 28198983]
43. Bergouignan A, Latouche C, Heywood S, Grace M, Reddy-Luthmoodoo M, Natoli A, et al. Frequent interruptions of sedentary time modulates contraction- and insulin-stimulated glucose uptake pathways in muscle: Ancillary analysis from randomized clinical trials. Sci Rep 2016 Aug 24;6:32044 [FREE Full text] [doi: 10.1038/srep32044] [Medline: 27554943]
44. Thyfault JP, Du M, Kraus WE, Levine JA, Booth FW. Physiology of sedentary behavior and its relationship to health outcomes. Med Sci Sports Exerc 2015;47(6):1301-1305. [doi: 10.1249/mss.0000000000000518]

Abbreviations

cpm: counts per minute
ERDF: European Regional Development Fund
FIQR: Revised Fibromyalgia Impact Questionnaire
HRQoL: health-related quality of life
MCS: mental component summary
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PAHELP: Physical Activity for Health Promotion
PCS: physical component summary
SF-36: 36-item Short-Form Health Survey
ST: sedentary time
UCEES: Unit of Excellence on Exercise and Health

Edited by G Eysenbach; submitted 06.05.19; peer-reviewed by M Miyachi, K Winfree, A Woodbury; comments to author 25.07.19; revised version received 01.09.19; accepted 24.09.19; published 19.03.20.

Please cite as:
Gavilán-Carrera B, Segura-Jiménez V, Acosta-Manzano P, Borges-Cosic M, Álvarez-Gallardo IC, Delgado-Fernández M
Patterns of Sedentary Time and Quality of Life in Women With Fibromyalgia: Cross-Sectional Study From the al-Ándalus Project
JMIR Mhealth Uhealth 2020;8(3):e14538
URL: http://mhealth.jmir.org/2020/3/e14538/
doi:10.2196/14538
PMID:32191211

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org